

**The dynamic behaviour of the chromosomal
protein SmcHD1 during X chromosome
inactivation**



Flavia Andreea Constantinescu

New College

University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy in Biochemistry

Abstract

Xist, the master regulator of X chromosome inactivation (XCI), employs a multitude of downstream factors to mediate chromosome-wide silencing. One of these factors is SmcHD1, which is a non-canonical member of the Smc family of chromosomal proteins. During XCI, SmcHD1 has been shown to play a role in gene repression, CpG islands methylation and remodelling of the chromosome architecture. Moreover, mutations in SmcHD1 have been associated with two genetic disorders: Facioscapulohumoral muscular dystrophy type 2 (FSHD2) and Bosma arrhinia and microphthalmia (BAMS). Nevertheless, the exact mechanism of action of SmcHD1 is not yet known in either physiological or pathological conditions.

In this thesis I aimed to uncover the molecular mechanism employed by SmcHD1 to perform its function by investigating its dynamic behaviour. In this respect, I setup a methodology to enable tracking of single molecules of SmcHD1 on the inactive X (Xi) chromosome, during *in vitro* monolayer (ML) differentiation of mouse embryonic stem cells. This method was then used to show how the dynamics of SmcHD1 change as silencing of the Xi chromosome progresses. Subsequently, since live imaging of SmcHD1 revealed a population-wide enrichment of the protein at pericentric regions, the role of SmcHD1's partner LRIF1 in this association was explored. Moreover, I observed that LRIF1 plays a role in the recruitment of SmcHD1 to the Xi in ML differentiation and, thus, investigated this further using the Xi-focused single-particle tracking methodology. Lastly, I have analysed the influence of three SmcHD1 mutations, including two disease-associated ones, on the dynamic behaviour of the protein as well as on its functionality in XCI. This has

allowed me to dissect the details of how SmcHD1 loads and unloads from the underlying chromatin, highlighting a key role for the ATPase activity of the protein in this process. Overall, the findings presented here will be key for understanding better the role of SmcHD1 in XCI and, further on, in the pathogenesis of FSHD2 and BAMS.

Acknowledgements

First, I would like to thank Neil for the opportunity to be part of his lab and the chance to work on such an interesting project. Thank you for the supervision and support throughout these four years and for giving me the freedom to explore my own ideas. Thank you for the suggestions and advice on this thesis.

I am extremely grateful to Adam for all his help, for taking the time to teach me all about microscopy and other techniques, for always answering any question I may have had and for always being patient, supportive and encouraging. Also, thank you for the comments on the thesis.

A special thank you to Tatyana and Mafalda for their scientific support and for their friendship. Also, thank you to Tatyana for making the lab a great place to work in. A big thank you also to Aleks for teaching me how to use the TIRF microscope, for contributing to this project and for his support. Thank you to Guifeng for interesting scientific discussions and for his contribution to this work. Also, many thanks to my second supervisor Stephan for useful discussions and suggestions. Thank you to all the past and present members of the Brockdorff lab for the support and thank you to everyone on the third floor of Biochemistry for creating an enjoyable atmosphere.

A massive thanks goes to all my friends for always being there for me and always managing to make me smile. I couldn't have done it without you. Thank you to Chiara, Joana, Agnese and Deborah for their support and all the fun times we shared. Thank you to my past and present housemates Yas, Clelia, Lucy, Kadi, Mev and Aleida for making the house feel like a home. Thank you to Ioio, Sandra and Maria for always keeping in touch despite the distance, for their understanding and support and for all the great adventures. A

special thank you to my friend, housemate and lab mate Clelia: we've been on this journey together from the very beginning and I am really grateful for it; it wouldn't have been the same without you!

Cel mai important, vreau sa multumesc familiei mele si in special parintilor mei pentru toata sustinerea lor in toti acesti ani, pentru ca sunt mereu alaturi de mine si ca ma incurajeaza. Nu as fi devenit persoana care sunt astazi si nu as fi putut sa imi indeplinesc visurile fara voi. Va multumesc din toata inima si va sunt profund recunoscatoare!

Contents

Abstract.....	1
List of figures.....	9
List of tables.....	12
List of abbreviations.....	13
Chapter 1: General Introduction.....	19
1.1 Chromatin and epigenetics.....	19
1.1.1 Chromatin structure and its functional implications.....	20
1.1.2 Euchromatin and Heterochromatin.....	23
1.1.2.1 Euchromatin.....	23
1.1.2.2 Heterochromatin	24
1.1.2.3 Constitutive heterochromatin.....	25
1.1.2.4 Facultative heterochromatin.....	26
1.1.3 3D chromatin organization.....	29
1.2 X chromosome inactivation.....	34
1.2.1 Xist RNA, regulation of its expression and functional repeats.....	35
1.2.2 Molecular mechanism of Xist-mediated X chromosome inactivation....	39
1.2.2.1 Xist spreading and localisation.....	39
1.2.2.2 Key roles for SPEN and the Polycomb pathway in establishing X chromosome silencing.....	42
1.2.2.3 Other players proposed to play a role in X chromosome silencing.....	45
1.2.2.4 Pathways associated with later stages and maintenance of XCI..	46

1.2.2.5 Architecture of the Xi chromosome.....	48
1.3 The non-canonical Smc protein SmcHD1.....	49
1.3.1 Structural and biochemical characteristics of SmcHD1.....	49
1.3.2 The role of SmcHD1 in X chromosome inactivation.....	54
1.3.2.1 Recruitment of SmcHD1 to the Xi chromosome.....	54
1.3.2.2 SmcHD1 and X-linked gene silencing.....	56
1.3.2.3 SmcHD1 and DNA methylation of Xi CGIs.....	58
1.3.2.4 The role of SmcHD1 in organising the architecture of the Xi...	59
1.3.2.5 SmcHD1 and imprinted XCI.....	62
1.3.3 The role of SmcHD1 in regulation of autosomal genes.....	63
1.3.4 SMCHD1 and human disease.....	64
1.4 Aims of this study.....	67
Chapter 2: Materials and methods.....	68
2.1 Plasmids and cloning.....	68
2.1.1 Homology vectors cloning for CRISPR-Cas9 targeting.....	68
2.1.2 sgRNA plasmids cloning for CRISPR-Cas9 targeting.....	72
2.2 Cell culture.....	72
2.2.1 Mouse embryonic stem cell growth.....	72
2.2.2 Generation of stable cell lines using CRISPR-Cas9-mediated homologous recombination.....	73
2.2.3 Monolayer differentiation.....	75
2.3 gDNA extraction and colony screening.....	76
2.4 Imaging methods.....	78
2.4.1 Single particle tracking.....	78

2.4.2. Single particle tracking analysis.....	79
2.4.3 Live cell imaging.....	80
2.4.4 Time course setup and analysis.....	81
2.4.5 FRAP.....	82
2.4.6 FRAP analysis.....	82
2.4.7 Immunofluorescence.....	83
2.4.8 H3K27me3 domain size analysis.....	84
2.5.1 Chromatin-associated RNA extraction and sequencing (chrRNA-seq).....	84
2.5.2 ChrRNA-seq analysis.....	85
2.6 Whole cell lysates, RIPA extracts, nuclear extracts and western blot.....	86
Chapter 3: Investigating dynamic behaviour of SmcHD1 on the inactive X chromosome using single-particle tracking.....	89
3.1 Introduction.....	89
3.2 Generation of a system for SmcHD1 single-particle tracking.....	91
3.3 Generation of an improved system which allows simultaneous visualisation of SmcHD1 and Xi domain in live imaging.....	95
3.4 Single-particle tracking to define SmcHD1 dynamics across the cell nucleus	100
3.5 Characterization of changes in SmcHD1 dynamics over the Xi territory during XCI.....	102
3.6 Discussion.....	106
Chapter 4: Studying the role of LRIF1 in modulating the dynamic behaviour of SmcHD1.....	110
4.1 Introduction.....	110
4.2 LRIF1 is required for localization of SmcHD1 to pericentric regions.....	112

4.3 LRIF1 is required for recruitment of SmcHD1 to the Xi chromosome during monolayer differentiation.....	116
4.4 The interplay between LRIF1-mediated and H2AK119ub1-dependent SmcHD1 recruitment pathways.....	130
4.5 Discussion.....	134
Chapter 5: Dissecting dynamic behaviour of SmcHD1 using SmcHD1 mutants	138
5.1 Introduction.....	138
5.2 Generation of a catalytically-dead, BAMS-associated and FSHD2-associated SmcHD1 mutant cell lines.....	140
5.3 Effects of SmcHD1 mutations on localization and dynamic behaviour over the whole nucleus.....	142
5.4 Effects of SmcHD1 mutations on recruitment and dynamic behaviour across the Xi chromosome.....	147
5.5 Effects of SmcHD1 mutations on its functionality during XCI.....	162
5.6 Discussion.....	165
Chapter 6: General Discussion.....	170
6.1 Investigating the molecular mechanism of SmcHD1 using live imaging.....	170
6.2 Two distinct localization pathways contribute to SmcHD1 enrichment over the Xi.....	171
6.3 The ATPase activity of SmcHD1 is required for its proper interaction with chromatin and functionality in XCI.....	174
Appendix.....	181
References.....	182

List of figures

1.1 Types of chromatin.....	21
1.2 The diverse mammalian Polycomb repressive complexes.....	27
1.3 Generation of Polycomb chromatin domain.....	29
1.4 3D organization of chromatin.....	31
1.5 Structure of tripartite cohesin ring.....	32
1.6 The mouse X inactivation centre.....	36
1.7 Xist RNA.....	38
1.8 Overview of XCI-associated events and chromatin modifications.....	39
1.9 Polycomb complexes recruitment cascade.....	44
1.10 SmcHD1 structure.....	51
3.1 Generation of a mESC line with endogenously Snap-tagged SmcHD1.....	94
3.2 Generation a mESC line with endogenously Halo-tagged SmcHD1 and eGFP-CIZ1 transgene expressed from the Rosa26 locus.....	98
3.3 SmcHD1 dynamics across the whole nucleus.....	101
3.4 SmcHD1 dynamics across the inactive X territory.....	104
4.1 Generation of an Lrif1 ^{-/-} mESC line.....	113
4.2 The effect of LRIF1 absence on SmcHD1's localization to pericentric regions...	114
4.3 The effect of LRIF1's absence on SmcHD1's dynamics across the whole nucleus.....	115
4.4 The effect of LRIF1 absence on SmcHD1 localization to the Xi chromosome.....	117
4.5 The effect of LRIF1 absence on SmcHD1 dynamics across the Xi chromosome...	119

4.6 The effect of LRIF1's absence on SmcHD1's dynamics assessed by FRAP.....	124
4.7 The effect of LRIF1's absence on X-linked gene silencing assessed by chrRNA-seq.....	127
4.8 Localization of LRIF1 over the Xi chromosome.....	129
4.9 Effect of MG132-induced H2AK119ub1 depletion on SmcHD1 localization to Xi.....	131
4.10 The effect of deletion of Xist PID region on SmcHD1 localization to Xi.....	133
5.1 Generation of SmcHD1 mutant mESC lines.....	141
5.2 The effect of SmcHD1 ^{E147A} mutation on the protein's localization to pericentric regions and dynamic behaviour across the whole nucleus.....	143
5.3 The effect of the SmcHD1 ^{G137E} mutation on the protein's localization to pericentric regions and dynamic across the whole nucleus.....	144
5.4 The effect of SmcHD1 ^{Y353C} mutation on the protein's localization to pericentric regions and dynamic behaviour across the whole nucleus.....	146
5.5 The effect of SmcHD1 ^{E147A} mutation on SmcHD1 localization to the Xi chromosome.....	148
5.6 The effect of SmcHD1 ^{E147A} dynamic behaviour of the protein across the Xi chromosome.....	150
5.7 The effect of SmcHD1 ^{G137E} mutation on Xi localization.....	151
5.8 The effect of SmcHD1 ^{G137E} mutation on the SmcHD1 dynamic behaviour across the Xi chromosome.....	153
5.9 The effect of SmcHD1 ^{Y353C} mutation on SmcHD1 localization to the Xi chromosome.....	155
5.10 The effect of SmcHD1 ^{Y353C} mutation on SmcHD1 dynamic behaviour across the Xi chromosome.....	157
5.11 The effect of SmcHD1 mutations on dynamic behaviour assessed by FRAP.....	160
5.12 The effect of SmcHD1 mutations on X-linked gene silencing assessed by chrRNA-seq.....	163

5.13 The effect of SmcHD1 mutations on Xi-associated H3K27me3 domain volume	165
6.1 Model of H3K9me3- and H2AK119ub1-dependent recruitment of SmcHD1 to Xi	174
6.2 A model for SmcHD1 interaction with chromatin and the Xi.....	178
A1 Histone H2B-Halo mESC line.....	181

List of tables

Table 1: Primers used for Gibson Assembly and site-directed mutagenesis..... 69

Table 2: Sequences of sgRNAs and associated PAMs used for CRISPR/Cas9..... 72

Table 3: Summary of cell lines derived in this study and associated transfection
and screening strategies..... 74

Table 4: List of primers used for screening of colonies from 96well plate..... 77

Table 5: Antibodies used for Western Blot..... 88

List of abbreviations

3D-SIM	Three-dimensional Structured Illumination Microscopy
4C-seq	Chromosome Conformation Capture-on-Chip Sequencing
AEBP2	AE Binding Protein 2
BAMS	Bosma Arrhinia Microphthalmia
bp	Base Pair
Cas9	CRISPR-associated protein 9
CBX	Chromobox Protein
CD	Chromatin Domain
Cdyl	Chromodomain Y-like Protein
CELF1	CUGBP Elav-Like Family Member 1
CGI	CpG Island
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation Sequencing
chrRNA-seq	Chromatin-associated RNA extraction and Sequencing
CIZ1	Cip1-Interacting Zinc finger protein
CpG	5'-C-phosphate-G-3' dinucleotides
cPRC1	Canonical PRC1
CTCF	CCCTC-binding Factor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DamID	DNA Adenine Methyltransferase Identification
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNMT	DNA methyltransferase

Dox	Doxycycline
DUX4	Double Homeobox 4
ECL	Enhanced Chemiluminescence
EED	Embryonic Ectoderm Development
eGFP	Enhanced Green Fluorescent Protein
ENU	<i>N</i> -ethyl- <i>N</i> -nitrosourea
EpiSC	Epiblast Stem Cell
ESC	Embryonic Stem Cells
E(var)	Enhancer of Variegation
EZH1	Enhancer of Zeste 1
EZH2	Enhancer of Zeste 2
FISH	Fluorescence In Situ Hybridization
FRAP	Fluorescence Recovery After Photobleaching
FSHD1/2	Facioscapulohumeral Muscular Dystrophy type 1/2
GFP	Green Fluorescent Protein
H2AK119ub1	Monoubiquitination of Lysine 119 of Histone H2A
H3K4me3	Trimethylation of Lysine 4 of Histone H3
H3K9me2/3	Di/Tri-methylation of Lysine 9 of Histone H3
H3K9ac	Acetylation of Lysine 9 of Histone H3
H3K27ac	Acetylation of Lysine 27 of Histone H3
H3K27me3	Trimethylation of Lysine 27 of Histone H3
H4ac	Acetylation of Histone H4
H4K20me1	Monomethylation of Lysine 20 of histone H4
HBiX1	HP1-binding Protein Enriched in Inactive X Chromosome 1

HDAC	Histone Deacetylase Complex
Hi-C	High-throughput Chromosome Conformation Capture
hnRNP	Heterogenous Nuclear Ribonucleoprotein Complex Protein
HP1	Heterochromatin Protein 1
Hsp90	Heat Shock Protein 90
HV	Homology Vector
IC	Interchromatin Compartment
iCLIP	Individual Nucleotide Crosslinking and Immunoprecipitation
ICR	Imprinting Control R6egion
IDR	Intrinsically Disordered Region
iXist-ChrX	Inducible Xist on Chromosome X
JARID2	Jumonji/ARID Domain-containing Protein 2
KDM2B	Lysine (K)-specific Demethylase 2B
KO	Knockout
LBR	Lamin B Receptor
LBS	LBR Binding Sites
LIF	Leukemia Inhibitory Factor
LINE	Long Interspersed Nuclear Element
lncRNA	Long Noncoding RNA
LRIF1	Ligand-dependent Nuclear Receptor-interacting Factor 1
m6A	N6-Methyladenosine
MATR3	Matrin 3
MBD-seq	Methyl-CpG Binding Domain Sequencing
MeCP2	Methyl CpG Binding Protein 2

MEF	Mouse Embryonic Fibroblast
mESC	Mouse Embryonic Stem Cell
METTL3/14	Methyltransferase-like 3/14
ML	Monolayer differentiation
MLL	Mixed Lineage Leukemia
Momme	Modifier of Murine Metastable Epialleles
NIPBL	Nipped-B-like Protein
NCoR	Nuclear Receptor Corepressor
NPC	Neural Progenitor Cells
NuRD	Nucleosome Remodelling and Deactylase
OCT4	Octamer-binding Transcription Factor 4
PAM	Protospacer Adjacent Motif
PAR-CLIP	Photoactivatable Ribonucleoside-enhanced Crosslinking and Immunoprecipitation
PBS	Phosphate Buffered Saline
Pcdh	Protocadherin
PcG	Polycomb Group
PCGF	Polycomb Group RING Finger
PEV	Position Effect Variegation
PHC	Polyhomeotic-like Protein
PID	Polycomb Interacting Domain
PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
PTBP1	Polypyrimidine Tract-binding Protein 1
PTM	Post Translational Modification

qPCR	Quantitative Polymerase Chain Reaction
RASER	Resolution After Single-stranded Exonuclease Resection
RBBP	Retinoblastoma Binding Protein
RBP	RNA-binding Protein
RBM15	RNA Binding Motif Protein 15
RE	Restriction Enzyme
RNA	Ribonucleic Acid
ROI	Region of Interest
RNAPII	RNA Polymerase II
RNF12	RING Finger Protein 12
RRM	RNA Recognition Motif domain
rTA	Reverse Tetracycline-controlled Transactivator
RT-qPCR	Reverse Transcription and Quantitative Polymerase Chain Reaction
RYBP	RING1 and YY1-binding Protein
SAFA	Scaffold Attachment Factor A
SAXS	Small-angle X-ray Scattering
Scc	Sister Chromatid Cohesion protein
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SEDTB1	SET Domain Bifurcated Histone Lysine Methyltransferase 1
sgRNA	Single Guide RNA
Smc	Structural Maintenance of Chromosomes
SmcHD1	Smc Hinge Domain-containing 1
SMRT	Silencing Mediator for Retinoid or Thyroid-hormone receptors
SNLP	STO mouse fibroblasts expressing Neomycin, Puromycin resistance and Lif genes

SNP	Single Nucleotide Polymorphisms
SOX2	SRY(Sex Determining Region Y)-box 2
SPEN	Split End
SPOC	SPEN Paralog and Ortholog
SPT	Single Particle Tracking
Su(var)	Suppressor of Variegation
SUZ12	Suppressor of Zeste 12
TAD	Topologically Associated Domain
TBP	TATA-Binding Protein
TDP-43	Transactive response DNA binding Protein 43 kDa
TetOP	Tetracycline-inducible Promoter
TF	Transcription Factor
TrxG	Trithorax Group
TTD	Tandem Tudor Domain
UBL	Ubiquitin-like Fold
vPRC1	Noncanonical Variant PRC1
WT	Wild-Type
WTAP	Wilms Tumor 1 Associated Protein
Xa	Active X Chromosome
XCI	X Chromosome Inactivation
Xi	Inactive X Chromosome
Xic	X Inactivation Centre
Xist	X-inactive Specific Transcript
YAF2	YY1 Associated Factor 2

Chapter 1: General Introduction

1.1. Chromatin and epigenetics

In eukaryotes, DNA is found within the membrane-enclosed organelle called the nucleus. While the average diameter of a nucleus is $\sim 2\text{-}10\mu\text{m}$, the strand of DNA has billions of base pairs (bp) and is $\sim 2\text{m}$ long. To fit inside the nucleus DNA has to be packaged in a highly organized structure termed chromatin. As well as aiding DNA packaging, chromatin organization aids the regulation of gene expression through epigenetic mechanisms, examples of which I will describe in more detail below. Epigenetics, a concept first introduced in 1942 by Conrad Waddington, refers to the study of changes in the transcriptional programme which take effect without altering the DNA sequence and are heritable through cell division (Waddington, 2012). Having a complex network of regulatory pathways, ready to switch genes on and off in a tightly controlled spatiotemporal manner, is highly important particularly during development: here, one cell, the fertilised zygote, needs to give rise to a multicellular organism which contains many different cell types, each having its own transcriptional programme. Epigenetic processes are characterized by plasticity, both at the initial stage of their setup, but, notably, also afterwards. They can be modulated both by genome intrinsic and environmental cues, which is, thus, key for ensuring adaptability of the organism. Importantly, some epigenetic pathways have been suggested to play a role in the mechanism of ageing and epigenetic alterations have been investigated in the context of several diseases such as obesity, neurological disorders and cancer. Overall, this highlights

the importance of further studying epigenetic processes. For a recent review of advances in the field and its importance see (Cavalli & Heard, 2019).

1.1.1 Chromatin structure and its functional implications

The term chromatin was first used by Walther Flemming in 1882 who saw a stainable material in the nucleus using light microscopy (the images reported are reproduced in (Paweletz, 2001)). Further studies, using X-ray crystallography, have shown the basic unit of chromatin is the nucleosome which represents 147bp of DNA wrapped left-handedly for ~1.65 turns around a histone octamer having two copies each of histone H2A, H2B, H3 and H4 (Kornberg, 1974; Luger et al., 1997). Between each pair of nucleosomes, there are ~50bp of free DNA, called linker DNA, which is bound by histone H1 to keep it in a more compact conformation when chromatin folds in higher-order structures (Figure 1.1 A).

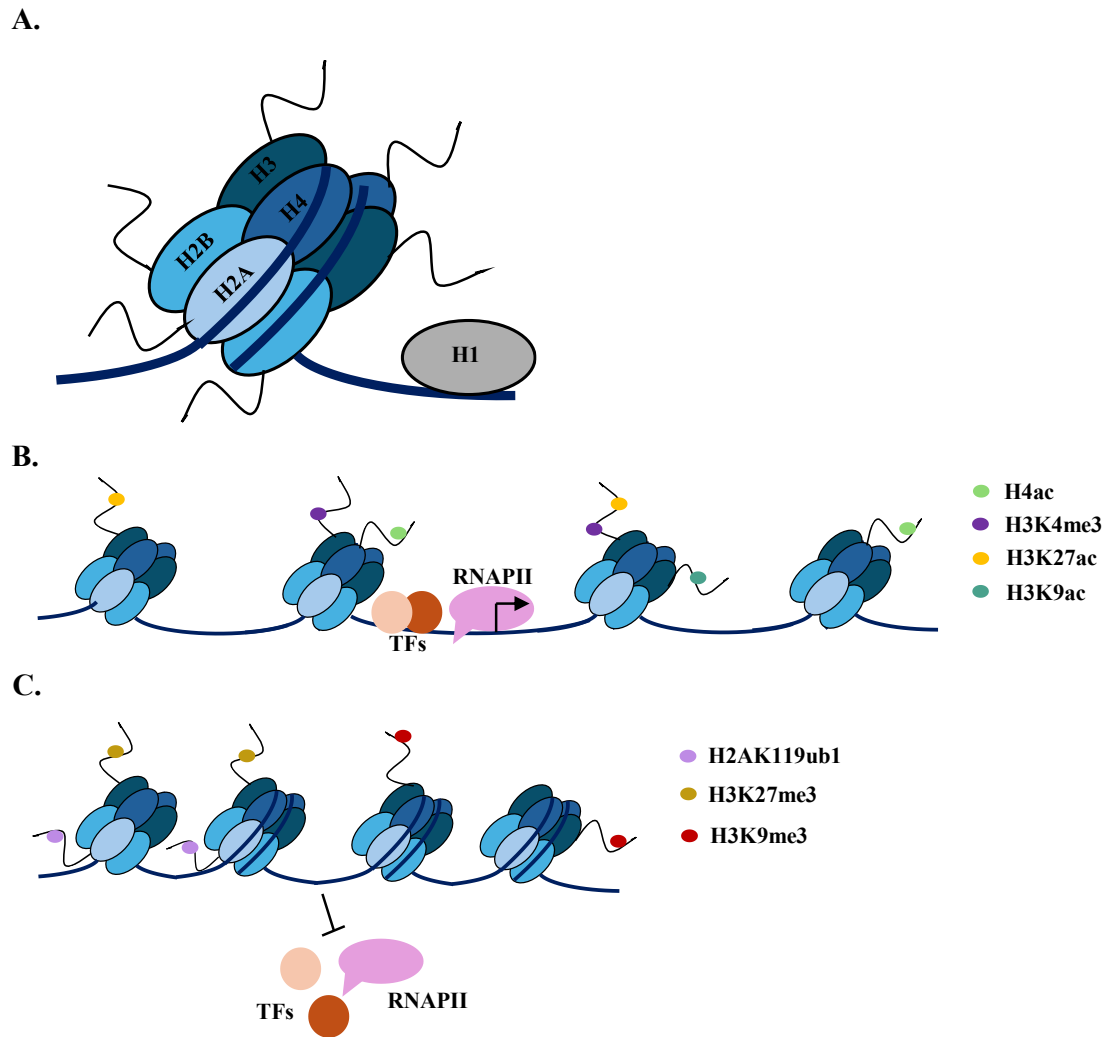


Figure 1.1 – Types of chromatin:

A. Schematic showing the structure of a nucleosome. A nucleosome contains two copies each of the histones H2A, H2B, H3 and H4. Histones have long flexible N-terminal tails which protrude from the nucleosome core and can be subject to post-translational modifications. Histone H1 binds the linker DNA between a pair of nucleosomes. **B.** Schematic illustrating an euchromatic region. The DNA is loosely packed, allowing access of the TFs and transcriptional machinery which will promote gene expression. Some examples of chromatin marks associated with euchromatin are H3K4me3 and acetylation of H3 and H4 (H3K27ac, H3K9ac, H4ac). **C.** Schematic showing characteristics of a heterochromatic region. The DNA is tightly packed, preventing access of the TFs and transcriptional machinery. There are two subtypes of heterochromatin: facultative which is associated with chromatin marks such as H2AK119ub1 and H3K27me3 and constitutive where H3K9me3 is found.

Notably, the packaging of DNA into nucleosomes can act as a regulatory barrier for gene expression by blocking the interaction of proteins such as transcription factors (TFs) and polymerases with DNA, due to them not being able to locate their target sequence while it is wrapped around the histone core. One mechanism to overcome this barrier is that cells employ pioneering factors. These have been defined as TFs which can recognize and bind their target sequence even when it is partially obscured and do so in a stable manner, before the binding of other TFs takes place (reviewed in (Zaret & Carroll, 2011)). Once bound, pioneering factors alter the local chromatin structure to make the DNA sequence accessible and, thus, stabilize their own binding as well as enable binding of other TFs. For example, the pioneering factor OCT4 recruits the ATP-dependent chromatin remodelling enzyme BRG1 and aids binding of other TFs, SOX2 and NANOG, ultimately aiding transcription of pluripotency-associated genes (King & Klose, 2017).

Other pathways that regulate chromatin accessibility make use of post-translational modifications (PTMs) of the long flexible N-terminal tails of histone proteins, which protrude from the highly ordered globular core of the nucleosome. The most common histone PTMs are lysine methylation (mono-, di- or tri-), lysine ubiquitylation, lysine acetylation and phosphorylation of serine and threonine sites. There are a multitude of enzymes which ensure deposition and removal of these marks, with a high level of specificity for the target residue (for a review on histone marks see (Bannister & Kouzarides, 2011)). Furthermore, the DNA itself can be methylated on the fifth carbon of a cytosine (5mC) when it is followed by a guanosine (CpG dinucleotide). The deposition is mediated by DNA methyltransferases (DNMTs) and can be reversed by TET methylcytosine dioxygenases (for a review on DNA methylation see (Greenberg & Bourc'h, 2019)).

Chromatin modifications can affect the regulation of transcription in several ways. Firstly, their presence can influence binding of TFs, either by enabling it, if they are readers

of that mark, or by deterring it, if their presence masks the DNA target recognition motif (see Section 1.1.2). Secondly, they can perturb the nucleosome's structure and, consequently, its interactions. A good illustration is given by histone acetylation which neutralizes the positive charge of the lysine amino group. This electrostatic change can destabilize the interaction between the histone octamer and DNA, which is negatively charged, as well as those between nucleosomes, required for chromatin folding into higher-order structures (Peppenella et al., 2014). Such changes lead to less condensed chromatin regions.

1.1.2 Euchromatin and Heterochromatin

Based on its characteristics, chromatin can be split into two types: euchromatin and heterochromatin, a division first proposed in 1928 by Emil Heitz. By studying the staining pattern of chromosomes in moss, he noted some parts of the chromosome remain stained through the cell cycle (heterochromatin), while others lose the staining after telophase (euchromatin) (summarised in (Passarge, 1979)).

1.1.2.1 Euchromatin

Euchromatin correlates with the parts of the genome which are relatively decondensed in interphase. These regions are gene-rich and characterized by an open and accessible chromatin structure that favours recruitment of TFs and the transcription machinery to drive gene expression (Figure 1.1 B). Some histone marks commonly found in such regions of active transcription are trimethylation of lysine 4 of histone H3 (H3K4me3) and acetylation

of histones H3 and H4. In humans, H3K4me3 is deposited by SET1A/B (SET domain-containing histone methyltransferase) or MLL (mixed lineage leukemia) enzymes (Cho et al., 2007; Dou et al., 2005; Issaeva et al., 2007; Lee & Skalnik, 2005; Lee et al., 2007) and it can be read and bound, amongst other factors, by the SAGA complex via the tandem Tudor domain (TTD) of one of its subunits (Vermeulen et al., 2010). SAGA is a histone H3-specific acetyltransferase complex and, as discussed above, this histone modification facilitates an open chromatin environment. H3K4me3 is also recognised by the TTD of the demethylase JMJD2A, which acts to remove trimethylation of lysine 9 of histone H3 (H3K9me3) that would otherwise favour recruitment of factors associated with transcriptional repression (Huang et al., 2006; Klose et al., 2006). Moreover, through their PHD finger domains, SET1 and MLL can bind their own histone mark and favour its spreading, further enforcing the open chromatin status (Eberl et al., 2013; Milne et al., 2010; Wang et al., 2010). Notably, there is a PHD domain also in the TAF3 subunit of the TFIID complex which is required for assembly of the pre-initiation complex required for RNA Polymerase II (RNAPII)-mediated transcription of protein-coding genes ((Vermeulen et al., 2007), reviewed in (Wozniak & Strahl, 2014)).

1.1.2.2 Heterochromatin

Heterochromatin is found at genomic locations which are compacted in interphase, gene-poor and associated with transcriptional repression (Figure 1.1 C). It is usually replicated later in S-phase compared to euchromatin (Gilbert, 2002) and is characterized by specific histone marks (see below). There are two further subtypes of heterochromatin: constitutive and facultative.

1.1.2.3 Constitutive heterochromatin

Constitutive heterochromatin is highly condensed throughout the entire lifetime of the organism and associated with tandemly repetitive DNA such as that found at centromeres and telomeres, and with transposable elements that give rise to large dispersed repeat families such as LINE repeats. Its function has been suggested to be as a genome stabilizer, preventing unwanted and potentially deleterious gene rearrangement between highly similar genetic sequences, by keeping them inaccessible (reviewed in (Janssen et al., 2018)). Important insights into constitutive heterochromatin were gained from studying position effect variegation (PEV). The PEV phenomenon takes place when a chromosomal rearrangement places a euchromatic gene in the vicinity of heterochromatin resulting in variegated expression due to the stochastic expansion of the heterochromatin domain. In 1930, Muller observed that an X-ray induced inversion moved the *white* gene, responsible for the pigmentation of the eye, close to pericentric heterochromatin and resulted in a patchy red/white eye, instead of the usual full red, in *Drosophila melanogaster* (Muller 1930). Further studies relied on generation of secondary mutations, using chemical mutagens, to identify modifiers of PEV, either suppressors (*Su(var)s*) or enhancers (*E(var)s*), including *Su(var)2-5* and *Su(var)3-9* (James & Elgin, 1986; Tschiersch et al., 1994), which encode two key factors for heterochromatin formation. *Su(var)3-9* encodes for a SET domain-containing histone methyltransferase which deposits H3K9me_{2/3}, while *Su(var)2-5* encodes for heterochromatin protein 1 (HP1) which binds H3K9me_{2/3} via its chromodomain (Jacobs & Khorasanizadeh, 2002; Schotta et al., 2002). Importantly, SU(VAR)3-9 and HP1 promote each other's binding to pericentric heterochromatin (Schotta et al., 2002). This reader-writer mechanism of heterochromatin spreading is conserved also in mammals. Furthermore, the cryo-EM structures of H3K9me₃ dinucleosomes complexed with either one of the three HP1

isoforms found in humans, HP1 $\alpha/\beta/\gamma$, has shown how HP1's ability to dimerize enables it to hold neighbouring nucleosomes closer together, further promoting chromatin compaction (Machida et al., 2018).

1.1.2.4 Facultative heterochromatin

Facultative heterochromatin occurs at genomic loci which can adopt either a condensed or decondensed chromatin organization, switching the transcriptional status of associated genes from repressed to active, depending on the stimuli to which the cell is exposed, for example, during differentiation and development (reviewed in (Trojer & Reinberg, 2007)). The Polycomb group (PcG) proteins are key factors in establishment and regulation of facultative heterochromatin. Their action is antagonized by the Trithorax group (TrxG) proteins. PcG proteins were first identified when studying *Drosophila* mutants suffering from homeotic transformations due to impaired expression of the Hox genes, which are responsible for correct segment specification along the anterior-posterior axis during embryo development (Ingham & Whittle 1980, (Lewis, 1978). The TrxG proteins correspond, in mammals, to the SET1A/B and MLL1-2 complexes, which, as discussed in Section 1.1.2.1, ensure deposition of the euchromatic mark H3K4me3.

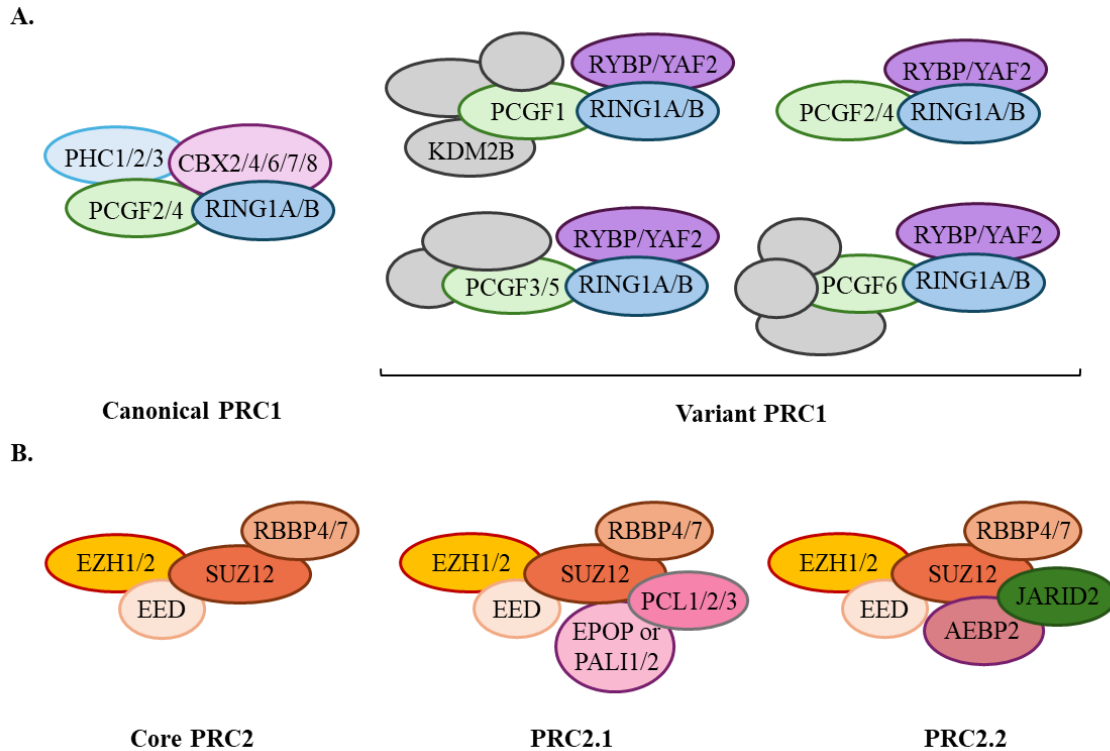


Figure 1.2 – The diverse mammalian Polycomb repressive complexes:

A. A canonical PRC1 forms around PCGF2/4 and includes one of the five CBX proteins, all of which have a chromodomain that can bind H3K27me3. A variant PRC1 can be formed around any of the PCGF1-6 and, depending on which one is included, associate with other auxiliary subunits such as KDM2B. RYBBP/YAF2 can bind to H2AK119ub1. Both cPRC1 and vPRC1s have RING1A/B as their catalytic core. **B.** In PRC2 complexes, EZH1/2 is responsible for the catalytic activity and EED ensures binding to H3K27me3. SUZ12 mediates interaction of core PRC2 with auxiliary subunits, leading to formation of PRC2.1 and PRC2.2. PCL1/2/3 can interact with non-methylated CpG dinucleotides, while JARID2 can bin H2AK119ub1.

Biochemical studies showed that most PcG proteins assemble into one of two major biochemical complexes, Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2). In mammals, the core of PRC1 is composed of one of the six Polycomb group RING finger (PCGF) proteins, PCGF1-6, and RING1A or RING1B, which are E3 ubiquitin ligases that ensure monoubiquitylation of lysine 119 of histone H2A (H2AK119ub1) (Wang et al., 2004). Based on the type of PCGF protein that they contain, these then assemble either with a CBX and a PHC subunit in a canonical PRC1 (cPRC1) or

with RYBP /YAF2 into a noncanonical variant PRC1 (vPRC1), as shown in Figure 1.2 A, with each complex having different catalytic potencies and binding capabilities. Core subunits of PRC2 are EED, SUZ12, RBBP4/7 and EZH2 (or its paralogue EZH1), the latter being a SET domain-containing histone methyltransferase which deposits trimethylation of lysine 27 of histone H3 (H3K27me3). Auxiliary factors associate with the complex to form two main types of complex, PRC2.1 and PRC2.2, thus modulating its catalytic activity and binding preferences (Figure 1.2 B). ChIP-seq analysis revealed that in vertebrates Polycomb complexes are mainly found at CpG islands (CGIs) (Ku et al., 2008), which are CpG-rich DNA segments of 1-2kb that remain non-methylated and are associated with about 70% of mammalian gene promoters (Deaton & Bird, 2011). For example, PCGF1-vPRC1 is recruited to CGIs thanks to the zinc-finger domain-containing protein, KDM2B, which specifically binds non-methylated CpG dinucleotides (Farcas et al., 2012).

Once recruited to their target location, Polycomb complexes form a so called repressive Polycomb chromatin domain through a series of feedback mechanisms, both PRC1/PRC2 intrinsic as well as between the two complexes, which are summarised in Figure 1.3 (reviewed in (Blackledge & Klose, 2021)). Early studies, notably using the *Drosophila* model, suggested that the Polycomb complexes follow a hierarchy where TF-mediated recruitment of PRC2 promotes deposition of H3K27me3, which is bound by cPRC1 that, in turn, deposits H2AK119ub1 and ultimately leads to gene repression. This was revised by more recent studies uncovering the existence of vPRC1 and showing that it can play a role upstream of PRC2 in recruitment to chromatin (Blackledge et al., 2014; Cooper et al., 2014; Cooper et al., 2016; Tavares et al., 2012).

A classical example of facultative heterochromatin is the inactive X (Xi) chromosome in cells of female mammals (reviewed in (Loda et al., 2022; Zyllicz & Heard,

2020)), which is at the forefront of this study and which will be discussed in detail in Section 1.2.

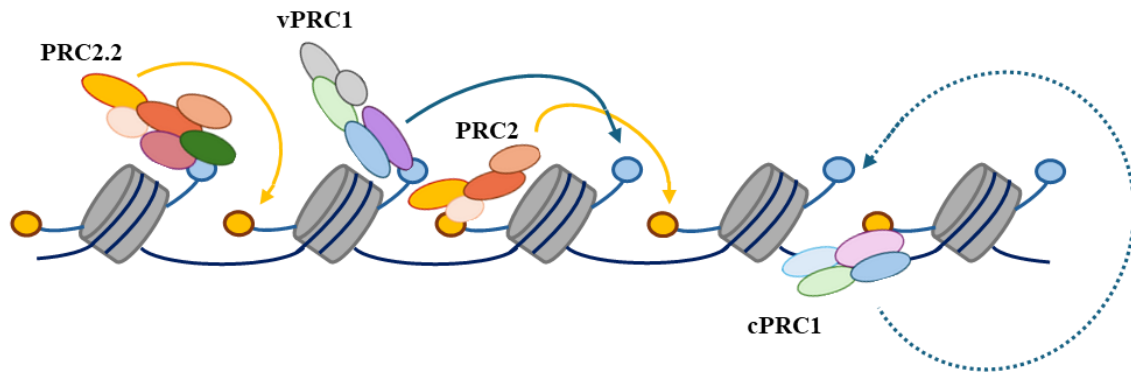


Figure 1.3 - Generation of Polycomb chromatin domain:

vPRC1s, via their RYBP/YAF2 subunit, can recognize their own histone mark, H2AK119ub1, promoting further vPRC1 binding and H2AK119ub1 deposition and, thus, spreading of the mark (intrinsic feedback mechanism). Similarly, PRC2, via its EED subunit, binds H3K27me3, and facilitates its spreading by reinforcing PRC2 binding and H3K27me3 deposition (intrinsic feedback mechanism). PRC1 and PRC2 can also sense and stimulate each other through their deposited histone modifications (inter-complexes feedback mechanism). cPRC1 binds H3K27me3, through its CBX subunit, and can deposit further H2AK119ub1, although it is less catalytically active than vPRC1s. PRC2.2 binds H2AK119ub1, via its JARID2-AEBP2 subunits, and leads to deposition of more H3K27me3. (adapted from (Blackledge & Klose, 2021))

1.1.3 3D chromatin organization

The chromatin organization principles and associated transcriptional consequences have so far been discussed from a 2D perspective, but, to have a comprehensive picture, it is important to consider also the higher-order 3D organizational levels of chromatin in interphase.

Insights into this level of organization have been offered by the development of the C techniques, such as high-throughput chromosome conformation capture (Hi-C) (Lieberman-Aiden et al., 2009), which couple proximity-based ligation and sequencing to determine which genomic loci interact. They showed that, over long distances (megabase pairs of DNA), similar types of chromatin interact with each other and, thus, segregate into two compartments: the A compartment representing transcriptionally active regions and the B compartment representing transcriptionally inactive regions (Lieberman-Aiden et al., 2009). Moreover, there are also interactions between certain genomic loci over shorter distances (hundreds of kilobase pairs of DNA) due to the folding of the chromatin fibre and regions where many such interactions occur give rise to topologically associated domains (TADs) which show up as squares on Hi-C maps (Figure 1.4 A-B) (Dixon et al., 2012). In parallel, microscopy approaches have also been used to understand better the 3D organization of the genome. Employing three-dimensional structured illumination microscopy (3D-SIM) coupled with 3D electron microscopy revealed chromatin is organized in a chain-like reticular network with nodes corresponding to nucleosome aggregations which were termed chromatin domains CDs, and which are separated by the interchromatin compartment (IC) space populated by RNA (Miron et al., 2020). Using resolution after single-stranded exonuclease resection – Fluorescence In Situ Hybridization (RASER-FISH), the authors observed that the spatial boundaries of an analysed TAD on the X chromosome overlapped with those of a CD, suggesting that CDs can correspond to TADs. Moreover, epigenetic markers associated with active transcription were seen to be localized at the interface between CDs and the IC, while those ensuring gene repression are found towards the core of the CD (Miron et al., 2020).

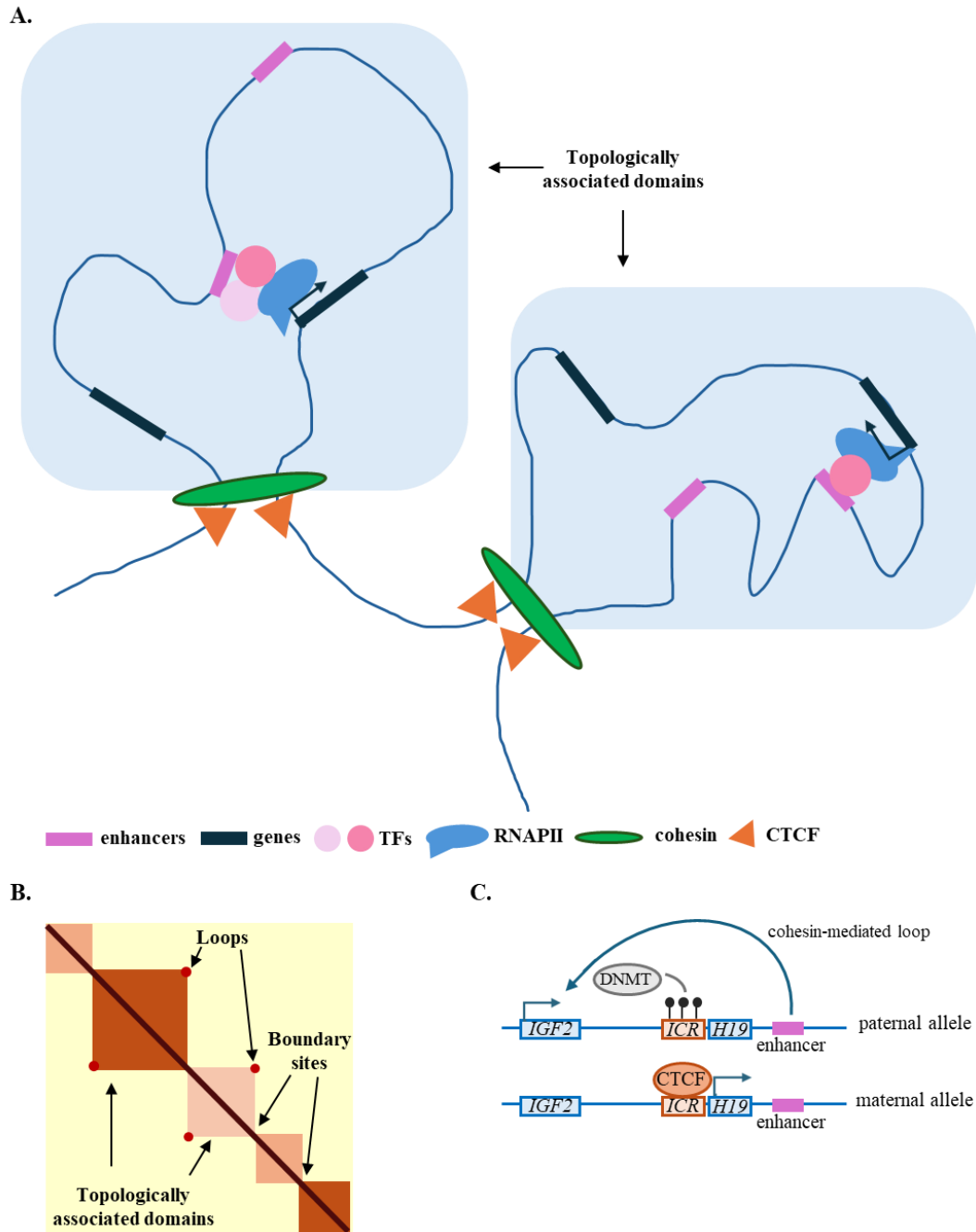


Figure 1.4 – 3D organization of chromatin:

A. Schematic of cohesin-CTCF mediated looping which leads to formation of TADs. Within the large loops, TFs and RNAPII mediate smaller loops. Ultimately, chromatin looping regulates promoter-enhancer interactions, which has a key impact on gene expression. **B.** Schematic of a Hi-C map showing how loops and TADs appear represented and the location of CTCF-bound boundary sites. (adapted from (Davidson & Peters, 2021)). **C.** At the *H19-IGF2* locus, the ICR is DNA methylated on the paternal allele, preventing CTCF binding at this site and allowing formation of a cohesin-CTCF mediated loop that enables interaction of a distant enhancer with the promoter of *IGF2* and, hence, its expression. The maternal ICR lacks DNA methylation, allowing CTCF binding and preventing formation of the afore-mentioned loop and, thus, *IGF2* expression.

A key player in higher-order 3D organisation of chromatin is the Structural maintenance of chromosomes (Smc) protein complex termed cohesin. The cohesin complex is ring-shaped and its structure is conserved across eukaryotes. In yeast, its core is composed of the two coiled-coil Smc proteins Smc1 and Smc3, which heterodimerise via their hinge domains and bind the α -kleisin subunit Scc1 through their nucleotide binding domains (Figure 1.5). Association of cohesin with DNA is tightly regulated by additional interacting subunits of the complex. For example, stable loading onto chromatin requires ATP-dependent engagement of the ATPase heads of Smc1 and Smc3 followed by ATP hydrolysis (Hu et al., 2011) and is mediated by Scc2-Scc4 (Ciosk et al., 2000), which also stimulates the ATPase activity (Petela et al., 2018).

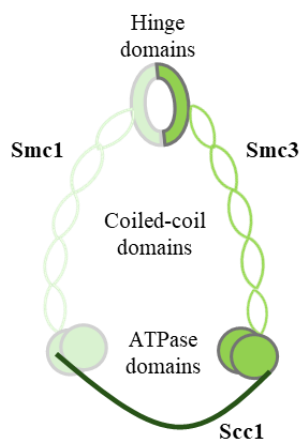


Figure 1.5 – Structure of tripartite cohesin ring

Based on Hi-C experiments, cohesin was proposed to organise the chromatin fibre through a loop-extrusion process, although its exact mechanism of action and regulation are still under debate (reviewed in (Corsi et al., 2023)). Advances in single-molecule microscopy approaches have recently enabled visualization of ATPase activity-dependent DNA loop formation by extrusion *in vitro*, by recombinant human cohesin in complex with NIPBL, the human homolog of Scc2 (Davidson et al., 2019; Kim et al., 2019; Tonkin et al., 2004). In mammalian cells, early ChIP data showed cohesin colocalizes with the DNA binding protein

CCCTC-binding factor (CTCF), with functional implications in gene expression (Parelho et al., 2008; Wendt et al., 2008). Later on, polymer simulations and genome-editing experiments were employed to explain the Hi-C maps of genome organization and it was proposed that loop extrusion by cohesin is stopped at boundary sites, which are represented by two convergent CTCF-bound sites, and leads to formation of the observed TADs (Fudenberg et al., 2016; Sanborn et al., 2015) (Figure 1.4 A-B). Recently, the crystal structure of a complex containing human SCC1 and CTCF revealed the molecular basis of their association and mutagenesis of the identified key interacting residues in CTCF resulted in complete loss of CTCF-anchored loops and TADs with less clear boundaries, although cohesin was still bound at CTCF sites, albeit at lower levels (Li et al., 2020). This suggests that CTCF fulfils a dual role of stopping extruding cohesin and stabilizing it at the base of the loop to prevent loop release (Li et al., 2020).

Such a 3D architecture of the chromatin fibre is important for the regulation of the transcriptional programme particularly when considering that expression of a gene is often regulated also by enhancers which act in *cis* and can be placed several hundred kilobases away from the promoter that they interact with. Thus, chromatin looping is key for bringing these genomic locations into close proximity as illustrated in Figure 1.4 A. It is important to note, that only some promoter-enhancer interactions, particularly those that take place across longer distances, are dependent on cohesin loop extrusion (reviewed in (Horsfield, 2023)), others rely on transcription and its associated regulatory factors such as Mediator (Hsieh et al., 2020; Lai et al., 2013). A classic example of a cohesin-CTCF mediated loop is depicted at the imprinted *H19-IGF2* locus, where cohesin was shown, using chromosome conformation capture and transcription assays, to be required for ensuring mono-allelic expression of *IGF2* as exemplified in Figure 1.4 C (Nativio et al., 2009). Notably, binding

of CTCF to the imprinting control region (ICR) is regulated by paternal-specific DNA methylation which takes place during spermatogenesis.

1.2 X chromosome inactivation

X chromosome inactivation (XCI) is the epigenetic process which takes place in cells of female mammals to ensure dosage compensation of X-linked genes between females, having 2 X chromosomes, and males, having only one X and a Y chromosome. A difference between female and male nuclei was first reported in 1949, upon staining of neurons with basic dyes which revealed a "nucleolar satellite" present only in female cells (Barr & Bertram, 1949). A functional interpretation of this observation was later on proposed by Mary Lyon who suggested that one of the two X chromosomes is randomly chosen to be inactivated early during embryonic development (Lyon, 1961).

XCI has been investigated in several mammalian species and, while some principles of action and/or players are conserved across them, the exact mechanisms of establishment and maintenance employed are quite diverse. Here, I will focus mainly on studies in mouse, a widely used model for the study of the XCI process. XCI studies in mouse have benefitted from access to key early-stage embryos when XCI is established and also XX mouse female embryonic stem cells (mESCs) in which the X inactivation process can be recapitulated in a tissue culture model.

In mouse development, XCI occurs in two waves. First, imprinted XCI takes place during preimplantation, starting from the 4-cell stage and leading to inactivation of the paternal X chromosome (Okamoto et al., 2004). In cells of the trophoctoderm and the primitive endoderm, the inactivation is permanent throughout development, while in the

inner cell mass of the blastocyst, the paternal X chromosome is reactivated (Mak et al., 2004; Okamoto et al., 2004; Takagi & Sasaki, 1975). Subsequently, from E5.5-E6.5 onwards, random XCI occurs in the embryo proper, leading to a mosaic inactivation of the maternal or paternal X chromosome across the cell population, and this status is then maintained across all further cell divisions (Mak et al., 2004). In the XX mESC model random XCI proceeds as mESCs are allowed to differentiate, for example into embryoid bodies (Rastan & Robertson, 1985).

1.2.1 Xist RNA, regulation of its expression and functional repeats

The master regulator of XCI has been identified, in both mice and human, as the long noncoding RNA (lncRNA) Xist (X-inactive specific transcript) which acts in *cis* (Brockdorff et al., 1991; Brockdorff et al., 1992; Brown et al., 1992; Brown, 1991). The 15-18kb long Xist transcript is localised in the nucleus, where it accumulates over the X chromosome from which it is transcribed, specifically the Xi chromosome, as shown using FISH assays (Brown et al., 1992; Clemson et al., 1996). Accordingly, Xist RNA is not translated and does not encode for any protein. Indeed, unlike protein-coding mRNAs, Xist RNA remains in the nucleus despite undergoing processing steps that are thought to promote the export of conventional mRNAs (Moore & Proudfoot, 2009). Critically, a gene knockout study creating a null mutation on one of the *Xist* alleles resulted in non-random XCI of the other X chromosome, showing that Xist is necessary in *cis* for XCI to occur (Penny et al., 1996). Moreover, expression of transgenic Xist was found to be sufficient to drive inactivation of an autosome on which it had been randomly integrated (Lee & Jaenisch, 1997).

The monoallelic upregulation of *Xist* is mediated by the X inactivation centre (Xic), a large region of around 800kb encompassing the *Xist* locus and which in mouse is organized into two TADs, one containing *Xist* and its positive regulators, TAD E, and one containing *Xist* negative regulators, TAD D (Figure 1.6) (Nora et al., 2012). This regulatory network is complex and still under investigation but is known to involve both *cis* and *trans* acting factors. Notably, initiation of XCI through *Xist* expression is tightly linked with onset of cellular differentiation and key pluripotency TFs, NANOG, OCT4, and SOX2, have been reported to contribute to downregulation of *Xist*, albeit their mechanism of action in this system is still unclear (Donohoe et al., 2009; Navarro et al., 2008; Schulz et al., 2014). Another important *Xist* regulator is RING finger protein 12 (RNF12) which is encoded ~500kb upstream of *Xist* and was initially shown to be a dose-dependent activator of *Xist* expression in that transgenic cells having extra copies of RNF12 presented inactivation of the X chromosome in male cells and inactivation of both X chromosomes in a proportion of female cells (Jonkers et al., 2009). Subsequently, it was shown that RNF12, which is a E3 ubiquitin ligase, targets degradation of the pluripotency TF REX1, that represses XCI initiation in a dose-dependent manner (Gontan et al., 2012).

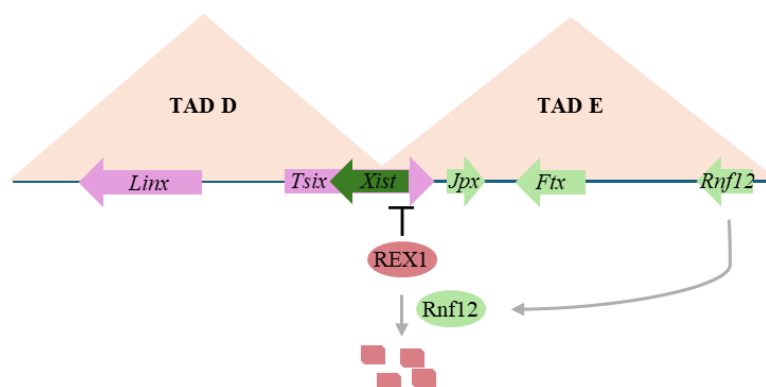


Figure 1.6 – The mouse X inactivation centre:

Schematic illustrating the organization of mouse Xic into two TADs, TAD E and TAD D, and some of the positive and negative, respectively, regulators that they contain.

Xist transcription is also controlled by several non-coding RNAs which are encoded within Xic. The most notable is the antisense transcript *Tsix* which acts as an *Xist* repressor (Lee et al., 1999). Other examples are the activators *Jpx* (Tian et al., 2010) and *Ftx* (Chureau et al., 2011) and the repressor *Linx* (Nora et al., 2012). The *Tsix* antisense transcription is initiated 15kb downstream of *Xist* and spans 40kb across the entire *Xist* locus. It is expressed biallelically before XCI is initiated, and then only from the active X chromosome allele transiently after *Xist* expression initiates (Lee et al., 1999). Heterozygous targeted deletion of a region encompassing the promoter of *Tsix* led to onset of non-random XCI with the mutated X chromosome becoming the repressed one (Lee & Lu, 1999). *Tsix* transcription has been proposed to exert its role in hindering *Xist* expression in *cis* by promoting chromatin modifications, such as DNA methylation, at the *Xist* promoter (Sado et al., 2006; Shibata & Lee, 2004; Sun et al., 2006). *Xist-Tsix* are located across the boundary separating the two TADs of Xic, with *Xist* promoter in TAD-E and *Tsix* promoter in TAD-D (Nora et al., 2012). Notably, a study engineered an inversion of the *Xist-Tsix* locus, exposing each promoter to the regulatory environment of the other one, and this caused ectopic *Xist* and *Tsix* expression, highlighting the key role of the bipartite TAD organization in mediating appropriate in *cis* transcriptional control (van Bemmelen et al., 2019).

Mono-allelic *Xist* transcription is the trigger for chromosome-wide silencing, achieved through the recruitment of key factors, including RNA-binding proteins (RBPs) that interact with specific sites on *Xist* RNA. Important elements for the recruitment of these factors are the six tandem repeat domains, A-F, of *Xist*, which are highly conserved across species, despite the low level of conservation of the primary sequence (Brockdorff et al., 1991; Brown, 1991; Nesterova et al., 2001) (Figure 1.7). Amongst the repeats, the most highly conserved is the 5' located A-repeat that was subsequently identified in a study looking at a large panel of transgenic *Xist* mutants, having various regions of it removed, as

being essential for the silencing function of Xist (Nesterova et al., 2001; Wutz et al., 2002). A key early study showed that Xist lacking the A-repeat is not able to trigger chromosome silencing, although localization over the chromosome was observed. Deletion of other tandem repeats, either alone or in combination, were found to disrupt Xist RNA localisation (Wutz et al., 2002).

A better understanding of the role of each of the different repeat domains of Xist, regarding its silencing function and localization, was made possible by the identification of the RBPs interacting with them, particularly by a series of independent studies which employed genetic screening assays or proteomic approaches (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015; Moindrot et al., 2015; Monfort et al., 2015). These RBPs and their role in XCI will be presented in detail in Section 1.2.2.

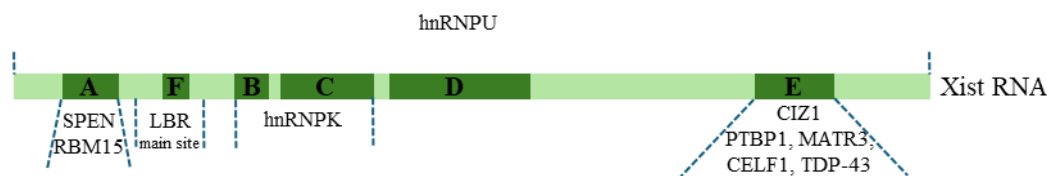


Figure 1.7 – Xist RNA:

Schematic illustrating the six repeat domains, A-F, of Xist RNA and the RNA-binding proteins associated with them.

1.2.2 Molecular mechanism of Xist-mediated X chromosome inactivation

Figure 1.8 illustrates an overview of the main events and chromatin modifications triggered by Xist in a stepwise manner in order to ultimately achieve chromosome-wide silencing of the chosen Xi.

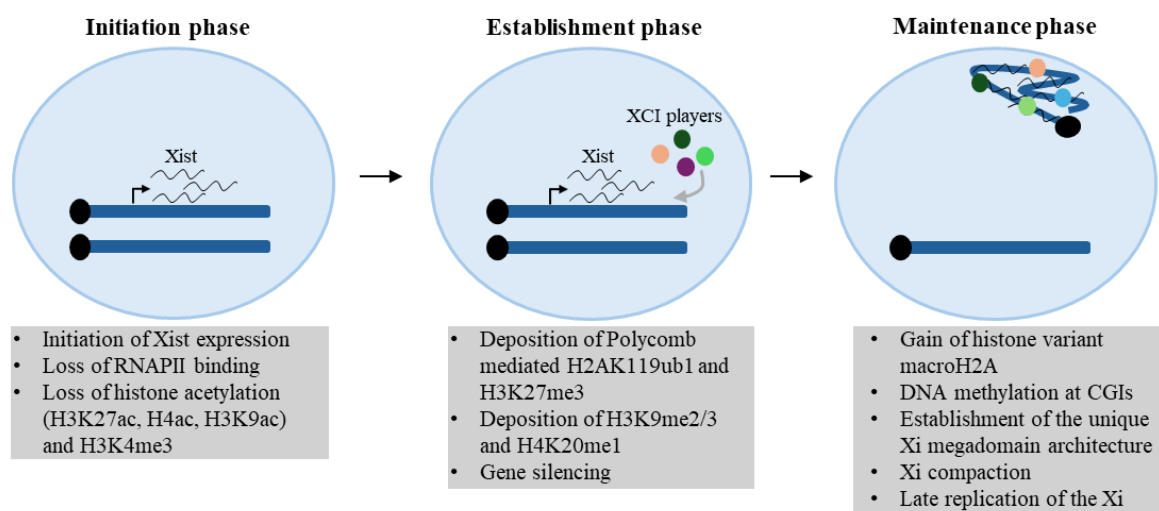


Figure 1.8 – Overview of XCI-associated events and chromatin modifications

1.2.2.1 Xist spreading and localisation

Two key aspects to consider are how Xist RNA molecules stay localized over the X chromosome chosen for inactivation and where and how many Xist molecules are required to facilitate chromosome-wide repression.

Regarding Xist localization, an early study reported that the RBP SAFA/hnRNPU has a role in ensuring *cis* localization of Xist over the Xi territory (Hasegawa et al., 2010). SAFA/hnRNPU binds to sites across the Xist transcript and based on its originally defined function in binding both DNA and RNA, is thought to function in anchoring Xist molecules with the Xi territory. However, the mechanism of SAFA/hnRNPU action in Xist localisation is not yet fully understood, and moreover its importance does not appear to be universal, for example in certain human cell lines (Hasegawa et al., 2010; Kolpa et al., 2016; Sakaguchi et al., 2016; Yamada et al., 2015).

Other RBPs involved in Xist RNA localisation interact with the Xist E-repeat element. The first characterized example was the nuclear matrix protein CIZ1 (Ridings-Figueroa et al., 2017; Sunwoo et al., 2017). Analysis of female mESCs taken through differentiation showed CIZ1 becomes enriched over the Xi territory immediately upon Xist expression, and that its association depends on the protein's C-terminal nuclear matrix anchor domain (Ridings-Figueroa et al., 2017). Female *Ciz1*-null mice show no defects in embryogenesis or early postnatal development, suggesting that CIZ1 is not required for the establishment of the repressed state of the Xi (Ridings-Figueroa et al., 2017). However, mouse embryonic fibroblasts (MEFs) derived from *Ciz1*-null mice or engineered to knockout *Ciz1* show dispersed Xist RNA across the entire cell nucleus (Ridings-Figueroa et al., 2017; Sunwoo et al., 2017). Moreover, adult *Ciz1*-null mice develop a lymphoproliferative disorder specifically in females and not in males, indicative of impaired XCI in lymphoid lineages (Ridings-Figueroa et al., 2017). A recent study confirmed, using 3D-SIM, that CIZ1 is necessary for ensuring Xi-limited localization of Xist in differentiated, but not in undifferentiated mESCs (Rodermund et al., 2021). This study confirmed that CIZ1 is a key player in the later, maintenance, stages of XCI. An independent study reported that CIZ1 and Xist are closely associated not only spatially, but also kinetically, with both showing a slow

recovery time over the Xi in fluorescence recovery after photobleaching (FRAP) assays, suggestive of a strong interaction with the E-repeat element (Markaki et al., 2021). Apart from CIZ1 the RBPs PTBP1, MATR3, CELF1 and TDP-43 have been reported to interact with the E-repeat element and to have a role in Xist RNA localisation (Pandya-Jones et al., 2020) .

Moving on to the number of Xist molecules and where on the chromosome they associate, the recent application of super-resolution imaging approaches indicated that there are surprisingly few, ~100-200 Xist molecules/cell (Cerase et al., 2014; Smeets et al., 2014; Sunwoo et al., 2015). As to where these molecules localise, initial studies relied on RNA crosslinking-based methods to identify the location of Xist across the Xi chromosome (Engreitz et al., 2013; Simon et al., 2013). This revealed that Xist initially accumulates at its transcription site, then transfers to preferential sites spread across the chromosome, which tend to be close to the *Xist* locus in 3D. Ultimately, Xist transcripts show a broad distribution over the entire domain (Engreitz et al., 2013). Xist seems to cover gene-rich regions first and gene-poor areas afterwards (Simon et al., 2013). The former is consistent with early work that reported that on metaphase chromosomes Xist localises in bands that correspond to gene-rich regions (Duthie et al., 1999).

More detailed insight into initial Xist localization came from a recent study employing 3D-SIM super-resolution microscopy to visualize individual Xist molecules at different timepoints following the onset of Xist expression (Rodermund et al., 2021). A two-step spreading mechanism was uncovered where an expansion phase, characterized by a radially enlarging Xist territory as the increasing number of RNA molecules move from the site of transcription to the periphery of the domain, is followed by a steady-state phase when there are less, but more stable Xist RNA molecules and the territory is no longer expanding (Rodermund et al., 2021). Notably, Xist was seen to undergo a process termed coupling,

where newly synthesized Xist molecules spread to sites where a presynthesized molecule is residing (Rodermund et al., 2021). The localization of Xist RNA molecules in pairs at about 50 sites across the Xi territory was reported also in another study relying on single-molecule super-resolution imaging (Markaki et al., 2021). Exactly how Xist molecules transfer to the preferred sites and, later on, across the entire chromosome remains to be determined.

1.2.2.2 Key roles for SPEN and the Polycomb pathway in establishing X chromosome silencing

The transcriptional repressor SPEN was identified in several independent studies as a key factor required for gene silencing during establishment of XCI (Chu et al., 2015; McHugh et al., 2015; Moindrot et al., 2015; Monfort et al., 2015). SPEN is an RBP which has four N-terminal RNA recognition motif domains (RRM1-4) and a C-terminal SPEN paralog and ortholog (SPOC) domain. SPEN was found to bind the A-repeat of Xist via RRM2-4 and its exact binding site has been mapped using individual nucleotide crosslinking and immunoprecipitation (iCLIP) (Chu et al., 2015; Lu et al., 2016; Monfort et al., 2015). Live imaging revealed SPEN associates with Xist immediately after the onset of the lncRNA's transcription and accumulates over the Xi forming a clearly distinguishable domain (Dossin et al., 2020). Two recent studies have reported a role for the previously uncharacterized central part of SPEN, a large intrinsically disordered region (IDR), in ensuring accumulation of SPEN over the Xi chromosome (Jachowicz et al., 2022; Markaki et al., 2021).

Removal of SPEN using a degron system or preventing its binding to Xist by engineering a deletion of the RRM domain (SPEN^{ΔRRM}) led to almost completely abolished gene silencing across the entire Xi in mESCs, 1 day after the onset of Xist expression (Dossin

et al., 2020; Nesterova et al., 2019). Importantly, SPEN was shown to be essential for XCI also *in vivo* as female *Spen*-knockout mouse embryos fail to undergo imprinted XCI (Dossin et al., 2020). The SPOC domain of SPEN is responsible for ensuring a large part of the protein's gene silencing function, as shown by comparison of silencing in a SPEN-SPOC mutant with a SPEN^{ΔRRM} mutant. However, SPEN also has SPOC-independent silencing functions (Bowness et al., 2022), and this is most likely linked to a requirement for SPEN for proper localization of Xist RNA over the Xi territory (Rodermund et al., 2021).

The SPOC domain mediates interaction with key downstream factors involved in XCI, most notably the corepressors NCoR/SMRT, which, in turn, bind and activate the histone deacetylase HDAC3 whose impairment also greatly affects X-linked gene silencing (Ariyoshi & Schwabe, 2003; McHugh et al., 2015; Zyllicz et al., 2019). Consistent with this link, H3K27ac and H4ac, are amongst the earliest chromatin modification events in XCI (Zyllicz et al., 2019). Proteomic analysis has shown that the SPOC domain also interacts with RNAPII subunits, the chromatin remodeller NuRD complex and the N6-methyladenosine (m6A) methyltransferase complex (Dossin et al., 2020). More studies are needed to uncover the importance of these interactions.

Besides SPEN, two factors that have long been known to be involved in XCI are the Polycomb complexes PRC1 and PRC2. Early studies reported their enrichment on Xi together with that of the chromatin PTMs that they catalyse, H2AK119ub1 and H3K27me3 respectively, at the early stages of initiation of XCI. It was further shown that recruitment of Polycomb depends on Xist expression (de Napoles et al., 2004; Fang et al., 2004; Kohlmaier et al., 2004; Plath et al., 2003; Silva et al., 2003). The pathway for recruitment was however debated for a long period due to conflicting evidence regarding whether they are direct Xist interactors and which Xist RNA elements are involved. Key studies have now revealed this mechanism (Almeida et al., 2020), which is illustrated in Figure 1.9. Specifically, the RBP

hnRNPK was identified as an interactor of Xist (Chu et al., 2015), binding to a region encompassing the B- and part of the C-repeat (Cirillo et al., 2016), which is essential for enrichment of both Polycomb complexes and associated histone modifications on the Xi (Pintacuda et al., 2017). Notably, artificial tethering of hnRNPK to an Xist transgene lacking the B/C-repeat rescues Polycomb recruitment (Pintacuda et al., 2017). hnRNPK associates with PCGF3/5-vPRC1 which in turn initiates the recruitment of PRC2 and other noncanonical PRC1 complexes (Almeida et al., 2017; Pintacuda et al., 2017).

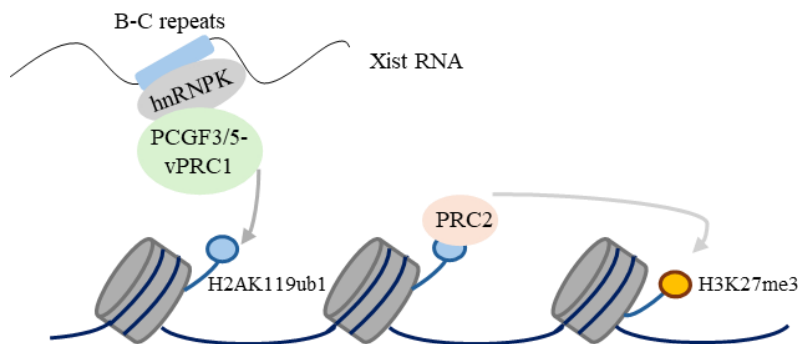


Figure 1.9 – Polycomb complexes Xi-recruitment cascade

The importance of the Polycomb pathway for XCI *in vivo* was highlighted by analysis of *Pcgf3-Pcgf5* double knockout embryos which displayed female-specific lethality (Almeida et al., 2017). Accordingly, various groups have reported defects in X-linked gene repression in mESCs models with perturbed functionality of the Polycomb pathway on the Xi (Almeida et al., 2017; Colognori et al., 2020; Nesterova et al., 2019; Pintacuda et al., 2017). It is important to note that the degree of silencing impairment varied across different experimental setups and further studies are needed to understand this better. A recent study has shed light on the interplay between the SPEN- and Polycomb-mediated silencing pathways showing that SPEN SPOC-dependent repression affects genes which are more

highly expressed in mESCs, while the SPEN SPOC-independent repression is mediated by the Polycomb system and, conversely, affects genes which are more lowly expressed (Bowness et al., 2022). Abrogation of both SPEN SPOC and Polycomb mediated repression resulted in complete loss of Xist-mediated gene silencing, suggesting that these two pathways function together for the establishment of chromosome silencing by Xist RNA (Bowness et al., 2022).

1.2.2.3 Other players proposed to play a role in X chromosome silencing

The identification of factors involved in the deposition of the post-translational RNA modification m⁶A, such as WTAP (regulatory subunit of the m⁶A methyltransferase complex) and RBM15 (recruiter of the complex), as interactors of Xist has led to the proposition that this modification may be involved in Xist-mediated XCI (Chu et al., 2015; McHugh et al., 2015; Moindrot et al., 2015). Indeed, Xist presents several sites where the m⁶A modification is present and knockdown of *Rbm15* and *Rbm15b* or of *Mettl3* (one of the two catalytic subunits of the complex together with METTL14) resulted in impaired silencing of some X-linked genes (Linder et al., 2015; Patil et al., 2016). However, subsequent studies employing mESCs where *Wtap*, *Rbm15* or *Mettl3* have been knocked out or the Xist m⁶A sites have been deleted showed only modest defects in Xi gene repression chromosome wide (Coker et al., 2020; Nesterova et al., 2019). More investigations will be crucial for uncovering the role of m⁶A machinery in XCI, perhaps one in Xist RNA stability (reviewed in (Wei, 2024)).

The Lamin B receptor (LBR), an inner nuclear membrane protein, has also been identified as an Xist interactor (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015)

with three main binding sites, termed LBR binding sites (LBSs), mapped across Xist (C. K. Chen et al., 2016). LBR was proposed to aid XCI by mediating association of the Xi chromosome to the nuclear lamina, which would, in turn, influence the 3D organization of the Xi allowing Xist to spread to and silence regions with active genes (C. K. Chen et al., 2016). However, analysis of the degree of gene silencing in *Lbr* null mESCs showed only a small impairment (Nesterova et al., 2019) in accordance with *Lbr* mutant mice showing no striking female-specific defects (Young et al., 2021).

1.2.2.4 Pathways associated with later stages and maintenance of XCI

Beyond the initial wave of Xist-mediated silencing, several downstream factors and epigenetic pathways are recruited as cellular differentiation progresses and these are thought to be important to reinforce and maintain the inactive status of the Xi and ensure its faithful propagation over the lifetime of the organism. Notably, in somatic cells, maintenance of XCI appears to be Xist-independent (Csankovszki et al., 1999). A crucial factor in mediating the final architecture of the Xi chromosome in the later phase of XCI establishment is the non-canonical Smc protein, SmcHD1 (Smc hinge domain-containing 1) and this will be presented in detail in Section 1.3.

Late-acting epigenetic changes on Xi include incorporation of the histone variant macroH2A (Mermoud et al., 1999), recruitment of the transcriptional corepressor Cdy1 in a H3K27me3- and H3K9me2-dependent manner (Escamilla-Del-Arenal et al., 2013) and DNA methylation of CGIs (Norris et al., 1991). As shown by analysis of differentiating female mESCs as well as female embryos, there are some CGIs which undergo methylation quickly after the onset of XCI, but the vast majority of them display slow methylation

dynamics (Gendrel et al., 2012). Notably, this main pathway is dependent on SmcHD1 (see Section 1.3.2.3) (Gendrel et al., 2012). Regardless of its timing, all CGI methylation on the Xi is mediated specifically by the *de novo* methyltransferase DNMT3B. Female embryos lacking *Dnmt3b* display impaired CGI methylation, unlike those lacking *Dnmt3a*, the other main *de novo* methyltransferase in mammals, or *Dnmt3l*, an accessory factor for DNMT3A/B which stimulates their catalytic activity (Gendrel et al., 2012). Female *Dnmt3b* null embryos are able to properly inactivate one X chromosome, showing that CGI methylation is dispensable for the establishment of XCI (Sado et al., 2004). However, independent studies using 5-azacytidine to inhibit DNA methylation have reported the treatment leads to reactivation of selected analysed genes on the Xi in female somatic cells (Graves, 1982; Mohandas et al., 1981). Of relevance here, combining deletion of *Xist* in somatic cells with 5-azacytidine treatment and with TSA treatment, which inhibits the action of histone deacetylases, leads to a synergistic increase in the percentage of cells displaying activation of a Xi-integrated reporter transgene compared to when only one or two of these modifications is introduced (Csankovszki et al., 2001). Overall, this suggests CGI methylation most likely acts redundantly with other mechanisms, to lock in the repressed state created by the early-acting XCI factors.

The proteins PTBP1, MATR3, CELF1 and TDP-43, which bind to the E-repeat of *Xist*, have been proposed to form condensates on the Xi thanks to homo- and heterotypic interactions and to play a role in maintaining X-linked gene repression, particularly once XCI becomes *Xist*-independent (Pandya-Jones et al., 2020).

1.2.2.5 Architecture of the Xi chromosome

The fully inactivated X chromosome is 1.2-fold more compacted and has a more spherical shape compared to its active counterpart which is flatter and elongated (Giorgetti et al., 2016; Teller et al., 2011). Chromosome conformation capture assays showed that Xi is characterized by absence of TADs and A/B compartments, which are present on the active X chromosome (Xa) and autosomes, and presents instead two mega domains separated by a boundary region which includes the *Dxz4* macrosatellite repeat (Deng et al., 2015; Giorgetti et al., 2016; Rao et al., 2014). The establishment of the Xi's unique architecture depends on the presence of intact Xist capable of carrying out gene silencing as an Xist mutant carrying an A-repeat deletion could not generate the *Dxz4* boundary (Giorgetti et al., 2016). The *Dxz4* array is DNA hypomethylated on the Xi and bound by CTCF (Chadwick, 2008; Horakova et al., 2012). Deletion of the boundary region causes extensive conformational changes leading to the fusion of the two mega domains, which highlights the key role of the region in ensuring the chromosome's architecture (Darrow et al., 2016; Giorgetti et al., 2016).

A recent study analysing in parallel the structural changes and gene silencing dynamics on the Xi during imprinted XCI has revealed that disappearance of the TADs occurs together with or after, but not before gene repression (Collombet et al., 2020). The Xi is characterized by reduced binding of the architectural organizers which promote TAD formation, cohesin and CTCF, compared to the Xa (Minajigi et al., 2015). Deletion of *Xist* in somatic cells causes the Xi to re-fold in an Xa-like conformation, characterized by TAD restoration mediated by cohesin, although gene reactivation was minimal probably thanks to the pathways of repression maintenance in place (Minajigi et al., 2015; Splinter et al., 2011). Notably, a few TAD-like structures are retained on the Xi and they are associated with

regions of open chromatin and residual transcription of a small subset of X-linked genes that escape XCI, termed escapees (Collombet et al., 2020; Giorgetti et al., 2016).

The mechanisms which mediate Xi compaction and its 3D architecture are not yet fully understood, but have been proposed to involve PRC1 and also SmcHD1 which will be discussed in Section 1.3.2.4.

1.3 The non-canonical Smc protein SmcHD1

1.3.1 Structural and biochemical characteristics of SmcHD1

A mutation in the previously uncharacterised *SmcHD1* gene was first mapped in a *N*-ethyl-*N*-nitrosourea (ENU) screen looking for modifiers of the variegated expression of a GFP transgene in male mice and was called *MommeD1* (Modifier of Murine Metastable Epialleles D1) (Blewitt et al., 2005). *MommeD1* is a point mutation that leads to appearance of a nonsense codon and is located in exon 23 of the 48 exon gene which encodes for a 2007 amino acid protein (Blewitt et al., 2008).

The domain organization of SmcHD1 classifies it as a noncanonical member of the Smc family of chromosomal proteins, with a comparison of the two types being shown in (Figure 1.10 A-B). Notably, the large size and flexible nature of the protein have hindered efforts to solve the crystal structure of full-length SmcHD1 so far, but studies looking at individual domains/regions have nevertheless shed light on its architecture and biochemical properties and, consequently, on how these may enable its functions. At the N-terminus, SmcHD1 presents a GHKL-type ATPase domain (Brideau et al., 2015; K. Chen, R. C. Dobson, et al., 2016; Pedersen et al., 2019), found in proteins such as DNA topoisomerase

II, the molecular chaperone Hsp90, the DNA-mismatch-repair enzyme MutL and histidine kinases. Presence of a GHKL ATPase contrasts with canonical Smc proteins which have a ABC-type ATPase domain that is split between the N- and C-termini in linear sequence with the two parts dimerising when the protein folds to form a functional domain (reviewed in (Hirano, 2006)). In a canonical Smc protein, the Smc hinge domain is centrally located, but, in SmcHD1, it is found near the C-terminus and, while the Smc hinge of SmcHD1 fold is similar to the one found in the canonical counterparts, it displays very little homology at the sequence level (Blewitt et al., 2008; Chen et al., 2020). The hinge domain of SmcHD1 is flanked by two α -helical coiled-coil forming domains, which are much shorter than those in canonical Smc proteins where they span the entirety of the regions between the hinge domain and the two ATPase heads and facilitate the characteristic fold of the protein via antiparallel coiled-coil interactions (reviewed in (Hirano, 2006)). Between the ATPase domain of SmcHD1 and the first α -helix, there is a long region whose structure and function are unknown, but structure prediction using AlphaFold suggests it is characterised by a series immunoglobulin like domains comprised of β -sheets (Dr. Tatyana Nesterova, personal communication). So far, only the crystal structures of the ATPase and hinge domains have been solved, while full-length recombinant SmcHD1 has been visualised using negative-stain electron microscopy (Brideau et al., 2015; Chen et al., 2020; Pedersen et al., 2019).

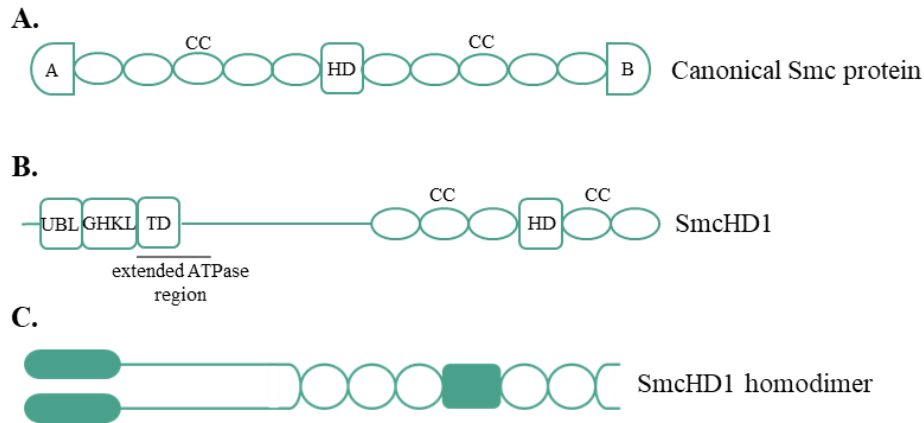


Figure 1.10 – SmcHD1 structure:

A. Domain composition of a canonical Smc protein: central hinge domain (HD), flanked by two long coiled-coil (CC) domains and a Walker A/B (A/B) motif on either end. **B.** Domain composition of SmcHD: the N-terminal ubiquitin like fold (UBL) precedes the GHKL ATPase domain and the transducer domain (TD). The extended ATPase region encompasses the TD and about 100 amino more downstream of it. At the C-terminus, there is a hinge domain flanked by two short coiled-coil domains. **C.** SmcHD1 dimerizes through its hinge domain.

The latter studies revealed a rod-like homodimeric molecule with a globular region on either end, most likely representing the ATPase and hinge domains (Brideau et al., 2015) (Figure 1.10 C). An analytic ultracentrifugation assay indicated recombinant SmcHD1 hinge domain dimerizes *in vitro* and, subsequently, immunoprecipitation showed the protein undergoes homodimerization, via its hinge domain, *in vivo* (Brideau et al., 2015). SAXS analysis of a construct containing the hinge domain and the flanking α -helical regions similarly reported homodimerization of the hinge and an intermolecular parallel coiled-coil arrangement between the pair of N-terminal α -helices and between the shorter C-terminal ones, which protrude from the globular hinge homodimer (K. Chen, P. E. Czabotar, et al., 2016).

Importantly, proteomic analysis did not identify any stoichiometric interacting proteins that could be analogous to the kleisin subunit found in cohesin's tripartite ring (see Section 1.1.3), suggesting that the SmcHD1 homodimer carries out core functions (Brideau

et al., 2015). Deletion of the hinge domain caused complete loss of chromatin-bound SmcHD1, as shown by a nuclear fractionation assay, indicating that either the hinge domain itself or protein dimerization is required for interaction with the underlying chromatin (Brideau et al., 2015). An independent study reported the hinge domain is able to bind DNA in a series of *in vitro* DNA binding analyses (Chen et al., 2015). Recently, the resolution of the hinge domain's structure allowed identification of positively charged patches on its surface, which were shown through mutagenesis of one amino acid at a time to be implicated in the nucleic acid binding property of SmcHD1 *in vitro*, although most of these mutations did not affect the nuclear localization pattern of the protein *in vivo*, possibly due to compensatory effects (Chen et al., 2020). Further characterisation will therefore be needed to understand the exact role of these positively charged residues in SmcHD1's association with DNA and, importantly, with chromatin.

Early studies on the ATPase domain of SmcHD1 relied on SAXS to analyse the structure of a recombinant SmcHD1 N-terminal region containing the ATPase domain and a section of about 300 amino acids downstream of it, termed the extended ATPase region and whose inclusion was required for ensuring the stability of the expressed product, and observed a similar conformation to that of full-length Hsp90 (K. Chen, R. C. Dobson, et al., 2016). Importantly, the domain was shown by two independent studies to be capable of ATP hydrolysis *in vitro* (Brideau et al., 2015; K. Chen, R. C. Dobson, et al., 2016). Mutation of the conserved glutamic acid (E147A) in Motif I of the of the ATP-binding pocket which essential for catalysis in other GHKL ATPases or treatment with radicicol, an inhibitor of GHKL ATPases, completely abolished or reduced the ATPase activity in a concentration-dependent manner, respectively (Brideau et al., 2015; K. Chen, R. C. Dobson, et al., 2016).

Recently, the crystal structure of the human SMCHD1 ATPase module has been solved in the presence of ATP, using a construct having the E147A mutation, and,

importantly, was found to undergo dimerization (Pedersen et al., 2019). As is the case for other GHKL ATPases, the module was found to contain a core GHKL ATPase catalytic domain (amino acids 110-395), which presents the characteristic Bergerat ATP-binding fold (an α/β sandwich with a long flexible loop termed the ATP lid) with four sequence motifs, and a transducer domain (amino acids 396-578), which has the fifth sequence motif. Moreover, a previously uncharacterized domain was identified, spanning amino acids 25-209 and adopting a ubiquitin-like fold (UBL). An UBL domain was not previously found in other GHKL ATPases. Notably, the transducer and UBL domains are required for dimerization of the N-terminal module, while the former is also needed for ATP hydrolysis, as suggested by gel filtration, cross-linking and ATPase assays using constructs lacking these regions (Pedersen et al., 2019). A subsequent study confirmed that a wild-type (WT) N-terminal SMCHD1 construct undergoes dimerization requiring the presence of the UBL domain and ATP binding (Gurzau et al., 2021). Notably, in human cells in which SMCHD1 is knocked-out, expressed constructs encoding for mouse SmcHD1 lacking the UBL domain or having the E147A mutation show altered nuclear distribution pattern, suggesting that both UBL-mediated dimerization and ATPase activity are required for correct chromatin localization of SmcHD1 (Gurzau et al., 2021). Overall, it seems that once the SmcHD1 homodimer, held together by the hinge domains with the aid of the two coiled-coil regions, binds ATP, it undergoes dimerization also at its N-terminal region. The implications of such conformational changes as well as the interplay with ATP hydrolysis for the function and interactions of the protein will be discussed in light of the findings of this project which will be presented in Chapters 3-6.

1.3.2 The role of SmcHD1 in X chromosome inactivation

Initial analysis of the *MommeD1* mutation showed that it was associated with increased expression of the GFP reporter transgene gene and of the retrotransposon-controlled *agouti viable yellow* (A^{vy}) metastable epiallele, suggesting the encoded protein functions as a transcriptional repressor (Blewitt et al., 2005). The latter effect was found only in female offsprings and, importantly, the mutation also led to homozygous female-specific midgestation embryonic lethality and hypomethylation of the CGI of the X-linked gene *Hprt*, suggesting SmcHD1 is involved in X chromosome silencing (Blewitt et al., 2005). Further analysis of embryonic and extraembryonic tissues from *SmcHD1^{MommeD1/MommeD1}* embryos at different developmental stages indicated that SmcHD1 plays a role in both random and imprinted XCI (Blewitt et al., 2008). Moreover, at E9.5/10.5, all X-linked analysed genes were found to present hypomethylated CGIs, while some of them displayed also defective gene repression and SmcHD1 was seen by immunofluorescence to be enriched over the Xi domain in MEFs (Blewitt et al., 2008). These initial observations set the scene for subsequent studies trying to understand better the role of SmcHD1 in XCI.

1.3.2.1 Recruitment of SmcHD1 to the Xi chromosome

Enrichment of SmcHD1 over the Xi territory is Xist-dependent in both mouse and human somatic cells (Jansz, Nesterova, et al., 2018; Nozawa et al., 2013; Wang et al., 2018). While the hinge domain of SmcHD1 can bind RNA oligos *in vitro*, it does so with no sequence specificity and, as shown by photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) in female primary neural stem cells and validation using

UV-RNA immunoprecipitation coupled with RT-qPCR, SmcHD1 does not directly bind any endogenous RNA genome-wide (Chen et al., 2015; Jansz, Nesterova, et al., 2018). Moreover, despite initially being identified as a potential interactor of Xist in a large proteomic study (Minajigi et al., 2015), subsequent validation using the same technique did not pull-down SmcHD1 together with Xist and high-resolution imaging of SmcHD1 and Xist revealed low overlap between their foci across the Xi domain, indicating that the two are not direct interactors (Jansz, Nesterova, et al., 2018).

Recruitment of SmcHD1 to the Xi depends on the B/C-repeat, hnRNP-K, and the PRC1 pathway: knock down of *hnRNP-K*, when XCI is in the maintenance phase, but also deletion of the *Xist* B/C-repeat in mESCs subsequently taken through differentiation, when XCI is being established, leads to loss of enrichment of SmcHD1 over the Xi territory (Jansz, Nesterova, et al., 2018). Notably, deletion of both *Ring1A* and *Ring1B* in MEFs impairs both localization of SmcHD1 at the Xi and its stability, while both these phenotypes are rescued by overexpression of WT Ring1B, but not a catalytic mutant. Moreover, super-resolution imaging showed higher co-localization of SmcHD1 and H2AK119ub1 foci than for any other potential recruiters previously investigated. Overall, these observations suggest that enrichment of SmcHD1 over the Xi chromosome depends on the PRC1-deposited H2AK119ub1 chromatin modification (Jansz, Nesterova, et al., 2018). The ATPase activity of the protein also seems to play a role in this localization as constructs encoding full-length SmcHD1 harbouring the E147A mutation expressed in *SmcHD1* null MEFs fail to accumulate at the Xi, although this mutation did not impact the protein's dimerization or general chromatin association (Brideau et al., 2015).

Regarding the timing of its recruitment, as previously mentioned, SmcHD1 is one of the later effectors in XCI. An early study employing female mESCs taken through embryoid body differentiation reported cells with SmcHD1 domains started to appear at day 7 and

extended across the whole population by day 9 (Gendrel et al., 2012). Recently, the timing of SmcHD1's accumulation over the Xi was re-evaluated using female mESCs engineered to have inducible Xist expression and differentiated towards neural progenitor cells (NPCs) (Bowness et al., 2022). Probably due to the more efficient expression of Xist in the latter cellular model, SmcHD1 enrichment over the Xi could be detected as early as day 3 and was present in almost all cells by day 5. Importantly, in the absence of differentiation SmcHD1 domains were not visible even after 10 days of Xist expression, indicating that recruitment to the Xi is dependent on a differentiation-associated factor/process, the details of which remain to be determined (Bowness et al., 2022).

1.3.2.2 SmcHD1 and X-linked gene silencing

As previously mentioned, early studies revealed female-specific embryonic lethality due to a loss of function mutation in *SmcHD1* and this was associated with impaired silencing of a few tested X-linked genes (Blewitt et al., 2008). Further chromosome-wide analysis using microarray expression profiling revealed about 10% of X-linked genes are upregulated in embryos lacking a functional SmcHD1 (Gendrel et al., 2013; Mould et al., 2013). Rather than being randomly distributed across the chromosome, the affected genes tend to associate across nine clusters which differ in terms of size and the number of upregulated genes contained within (Gendrel et al., 2013).

Given its late recruitment to the Xi, SmcHD1 was initially proposed to play a role in the maintenance phase of XCI. Higher sensitivity transcriptional analysis in MEFs derived from SmcHD1 *null* embryos or in NPC clones differentiated from SmcHD1 knockout mESCs, which importantly allow for distinction of the Xi versus the Xa allele based on

single nucleotide polymorphisms (SNPs), showed around half of X-linked genes are upregulated to various levels (Gdula et al., 2019; Sakakibara et al., 2018; Wang et al., 2018). Gene expression was accompanied by CGI hypomethylation as well as depletion of the heterochromatic mark H3K27me3 in associated areas, a feature uncovered by ChIP-seq, as immunostaining for this chromatin modification did not show any obvious changes in its enrichment over the Xi (Blewitt et al., 2008; Gdula et al., 2019; Sakakibara et al., 2018; Wang et al., 2018). Interestingly, the derepressed genes present lower enrichment of H3K27me3 at the blastocyst stage, suggesting functional SmcHD1 is needed to mediate further deposition of the mark at these sites and achieve steady silencing (Sakakibara et al., 2018). Notably, knocking out *SmcHD1* in MEFs, when XCI is, thus, completely established, does not lead to transcriptional activation of the X-linked genes, which is consistent with SmcHD1 being required to strengthen the inactive status of the Xi during XCI establishment, and then becoming dispensable, probably due to redundancy with other XCI maintenance mechanisms as discussed above (Gdula et al., 2019; Sakakibara et al., 2018).

Indeed, a recent study has addressed the role of SmcHD1 in the earlier days of XCI, using mESCs with inducible Xist differentiating towards NPCs (Bowness et al., 2022). Using the classification of SmcHD1-dependency based on the earlier study in embryonic *SmcHD1* null derived MEFs, the authors observed that SmcHD1-dependent genes are generally characterized by slower silencing kinetics in NPC differentiation and correspond to genes which fail to achieve full repression in undifferentiated cells that have expressed Xist for several days (Bowness et al., 2022; Gdula et al., 2019). Importantly, SmcHD1 knockout mESCs start to display a plateau in silencing of SmcHD1-dependent and partially-dependent genes from day 5 of NPC differentiation onwards, when SmcHD1 is enriched over the Xi population-wide, while SmcHD1-independent genes undergo complete inactivation (Bowness et al., 2022). Therefore, after the initial wave of silencing mediated

by the SPEN and Polycomb pathways, as described above, SmcHD1 is recruited to the Xi and is crucial for finalising chromosome-wide silencing (Bowness et al., 2022).

1.3.2.3 SmcHD1 and DNA methylation of Xi CGIs

As mentioned earlier, in the first studies on SmcHD1, *SmcHD1^{MommeD1/MommeD1}* female embryos displayed hypomethylation of all 16 analysed X-linked CGIs (Blewitt et al., 2008; Blewitt et al., 2005). Subsequent MBD-seq analysis in MEFs derived from *SmcHD1* null embryos showed that CpG methylation of late methylating CGIs, which is the predominant category on Xi CGIs, depends on SmcHD1 (Gendrel et al., 2012). Moreover, fast methylating CGIs can achieve significant methylation independently of SmcHD1, but at a lower level compared to when the protein is present, suggesting SmcHD1 is involved in DNA methylation also at these CGIs (Gendrel et al., 2012). In line with such a function, the hinge domain of SmcHD1 was shown to bind methylated DNA oligos *in vitro* (Chen et al., 2015), although exactly how this association would mechanistically aid SmcHD1 remains to be determined.

Thus, given the extensive role of SmcHD1 in CGI methylation, it was tempting to speculate that this is the pathway through which the protein promotes gene silencing. However, analysis of expression level by qPCR showed that *Dnmt3b* null female embryos have modest or no upregulation of X-linked genes derepressed in *SmcHD1* null female embryos, indicating that CGI hypomethylation is not the sole reason behind gene activation in cells undergoing XCI without functional SmcHD1 (Gendrel et al., 2013).

1.3.2.4 The role of SmcHD1 in organising the architecture of the Xi

SMCHD1 was first proposed to play a role in mediating chromosome architecture of human Xi (Nozawa et al., 2013). Proteomic analysis showed SMCHD1 associates with HBiX1 (HP1-binding protein enriched in inactive X chromosome 1), an interactor of HP1 proteins whose mouse homolog, LRIF1, was later found to also bind SmcHD1 (Brideau et al., 2015; Nozawa et al., 2013). Knockdown of SMCHD1 or HBiX1 in somatic human female cells led to a less compacted Xi, more similar to the Xa, as assessed by FISH, and, more specifically, to less compacted Xi-associated foci of H3K9me₃, which is normally enriched across human Xi and to which HBiX1 binds via HP1 (Nozawa et al., 2013). In contrast, the distribution of H3K27me₃ foci was not affected by the reduced levels of SMCHD1/HBiX1. As expected, knockdown of Xist, which impairs enrichment of SMCHD1 over the Xi, also causes decompaction of the Xi, but reduced levels of SUZ12 do not cause such changes. Deconvolution microscopy showed H3K9me₃ and H3K27me₃ occupy distinct regions of the Xi and SMCHD1 and HBiX1 are located at the interface between these two domains. Knockdown and mutant analyses revealed that SMCHD1 associates indirectly with the H3K9me₃ regions via HBiX1, while, conversely, HBiX1 requires SMCHD1 to localize over the H3K27me₃ regions, although how SMCHD1's own, probably indirect, association with H3K27me₃ is mediated is not known. Thus, the interaction between SMCHD1 and HBiX1 contributes to maintenance of Xi compaction by bringing the H3K9me₃ and H3K27me₃ regions closer together, independently of PRC2-H3K27me₃ (Nozawa et al., 2013). A subsequent analysis in mouse somatic cells found that *Lrif1* is not required to localise SmcHD1 to Xi (Brideau et al., 2015). Of relevance here is that the strong H3K9me₃ enrichment seen on the human Xi is not apparent on the mouse Xi (Rens et al., 2010). Potentially linked to this difference between Xi in mouse and human, deletion of

SmcHD1 in female primary neural stem cells did not alter the volume of the Xi, indicating that, in mice, SmcHD1 does not play a role in maintaining Xi compaction (Jansz, Keniry, et al., 2018).

Subsequently, studies have relied on the use of the chromosome conformation capture technique Hi-C to study the role of SmcHD1 in the architecture of mouse Xi (Gdula et al., 2019; Jansz, Keniry, et al., 2018; Wang, Colognori, et al., 2019; Wang et al., 2018). *SmcHD1* loss of function mutations both pre- and post-XCI establishment attenuate the characteristic megadomain structures of the Xi, causing an increase in short-range interactions and restoration/strengthening of TADs and compartments, leading to an overall Xa-like architecture (Gdula et al., 2019; Jansz, Keniry, et al., 2018; Wang, Colognori, et al., 2019; Wang et al., 2018). Notably, in a germline mutant, the degree of TAD re-establishment is correlated with increase in transcriptional activity, decrease in H3K27me3 and increase in CpG methylation, while in somatic mutants, as mentioned before, there is no gene upregulation, highlighting that, while chromosome organisation plays a role in the establishment of gene repression, it becomes dispensable during the maintenance of XCI, due to redundancy with other pathways (Gdula et al., 2019; Jansz, Keniry, et al., 2018). Indeed, only when post-XCI SmcHD1 depletion was coupled with inhibition of DNA methylation via 5-azacytidine treatment, could gene reactivation be observed, particularly affecting genes whose initial silencing is dependent on SmcHD1. This experiment indicated that SmcHD1 and DNA methylation act synergistically to protect the inactive status of the Xi (Wang, Colognori, et al., 2019).

The compartments observed in the absence of functional SmcHD1 were reported by one study to be distinct from the usual A/B-compartments found on the Xa and autosomes and were termed S1 and S2, with S1 being more gene- and Xist-rich (Wang et al., 2018). S1/S2 compartments were observed not only in differentiated cells obtained from mESCs

lacking SmcHD1, but also at intermediate timepoints during differentiation of WT mESCs, before SmcHD1's recruitment to the Xi. As mentioned in Section 1.2.2.1, Xist tends to spread first to gene-rich and then to gene-poor region and, thus, it was proposed, given SmcHD1's broad coverage across the entire Xi as seen using DamID, that SmcHD1 bridges the S1 and S2 compartments facilitating transfer of Xist from the S1 to S2 regions during establishment of XCI (Wang et al., 2018). In a subsequent study, it was also shown SmcHD1 is required to preserve the distribution of Xist across both compartments during the maintenance phase (Wang, Colognori, et al., 2019). Moreover, post-XCI deletion of Xist or depletion of PRC1 or its recruiter hnRNPK in either WT or somatic *SmcHD1* mutant cells, impairs appearance of the S1/S2 compartments, which led the authors to propose that, following its Xist-dependent recruitment, PRC1 facilitates formation of these compartments, potentially through phase separation (Wang, Colognori, et al., 2019).

The observed chromosome architecture changes in the absence of SmcHD1 are accompanied by restoration of CTCF and cohesin binding, suggesting SmcHD1 may facilitate establishment of the unique architecture of the Xi by preventing association of these architectural proteins which, as discussed in Section 1.1.3, normally mediate 3D chromatin organization (Gdula et al., 2019; Wang et al., 2018). Notably, such an antagonistic relationship between SmcHD1 and CTCF binding, with consequences for gene expression, is not exclusive to the Xi, having been previously observed also at an autosomal target of SmcHD1, the protocadherin (*Pcdh*) genes (Chen et al., 2015).

Notably, the architectural remodelling caused by loss of function mutations of SmcHD1, either germline or somatic, also perturbed the asynchronous replication of the Xi, anticipating it towards early-mid S phase (Gdula et al., 2019; Nozawa et al., 2013; Poonperm et al., 2023; Sakakibara et al., 2018). The entire Xi is normally replicated mid/late during the S phase, unlike the Xa and other autosomes whose gene-rich and gene-poor regions are

replicated early or late, respectively, during S phase (Aladjem & Fu, 2014). Importantly, not all regions across the Xi are affected to the same extent, with those having the most advanced replication protruding out of the Xi core, as seen by 4C-seq and DNA-FISH, and being correlated with gene activation and associated epigenetic modification such as loss of H3K27me3 (Gdula et al., 2019; Poonperm et al., 2023).

The relationship between SmcHD1's architectural and gene silencing roles is currently not well understood, but a recent study has suggested the two functions are independent of each other (Tapia Del Fierro et al., 2023). The authors investigated a mutation of SmcHD1 called *MommeD43* (A667E) which was found in the same ENU mutagenesis screen that uncovered the *MommeD1* mutation (Blewitt et al., 2005; Tapia Del Fierro et al., 2023). *MommeD43* is placed in the extended ATPase domain of SmcHD1, but it does not affect the ATPase activity or conformation *in vitro* of a recombinant N-terminal construct (Tapia Del Fierro et al., 2023). Interestingly, the mutation caused chromatin conformation changes that are highly correlated with those observed in the absence of SmcHD1, but these were not accompanied by impaired gene silencing, rather by increased repression of some of SmcHD1's autosomal targets such as the *Hox* cluster (see Section 1.3.3). This indicates the long-range chromatin interactions mediated by SmcHD1 are not directly involved in the silencing of its target genes (Tapia Del Fierro et al., 2023).

1.3.2.5 SmcHD1 and imprinted XCI

The exclusive paternal expression of Xist during imprinted XCI in mice is ensured by deposition of a repressive non-canonical Polycomb-mediated imprint over the Xist locus on the maternal allele, in the oocyte (Inoue et al., 2017; Mei et al., 2021). Maternal deletion of

SmcHD1 was seen to cause upregulation of maternal Xist in a subset of both female and male E2.75 embryos, with the former displaying biallelic expression in the same cells, as seen by RNA-FISH (Wanigasuriya et al., 2022). Inappropriate Xist expression caused partial silencing of the maternal X in male embryos and slightly impaired silencing of the paternal X in female embryos. By E3.5, the normal pattern of Xist expression is however reinstated in the outer trophoctoderm layer in both maternal SmcHD1-deficient female and male embryos. This suggests that the absence of maternal SmcHD1 only affected imprinted expression of Xist, but not initial deposition and maintenance of the Polycomb-mediated imprint whose function can probably be exerted through zygotic SmcHD1 which is expressed from this timepoint onwards (Wanigasuriya et al., 2022). However, even though Xist expression is normal, in female E14.5 placentae of maternal SmcHD1-deficient embryos, about half of the analysable X-linked genes were upregulated and the CGIs displayed hypomethylation. The effect on gene silencing was more severe than in a case of *SmcHD1* haploinsufficiency (Wanigasuriya et al., 2022). Overall, this suggests maternal SmcHD1 acts downstream of the Polycomb complexes and is crucial for establishment of an epigenetic landscape which is needed to ensure gene silencing later on during imprinted XCI (Wanigasuriya et al., 2022).

1.3.3 The role of SmcHD1 in regulation of autosomal genes

Apart from its function in XCI, SmcHD1 has a role in the silencing of selected autosomal loci, typically found in gene clusters such as the *Pcdh*, *Hox*, *Snrpn*, *Igf2r*, with notably most of the regulated genes normally being monoallelically expressed (Chen et al., 2015; Gendrel et al., 2013; Jansz, Keniry, et al., 2018; Mould et al., 2013). Gene upregulation at the *Snrpn*

and *Pcdh* clusters is associated with CGI hypomethylation, which for the *Snrpn* genes leads to loss of imprinting (Gendrel et al., 2013; Mould et al., 2013), and with changes in chromatin interactions at the *Pcdh* and *Hox* clusters (Jansz, Keniry, et al., 2018). For example, Hi-C analysis following *SmcHD1* deletion in neural stem cells, showed the *HoxB* cluster lost long-range interactions and strengthened short-range interactions, and displayed posterior Hox gene activation not only in a cellular model, but also *in vivo*, with *SmcHD1* null embryos displaying defects in the morphology of their vertebral column (Jansz, Keniry, et al., 2018). Altogether, these observations suggest that SmcHD1 employs similar pathways in the silencing of its autosomal targets as it does for the X-linked genes, although exactly how SmcHD1 fulfils these functions to ultimately achieve gene repression is not yet understood.

1.3.4 SMCHD1 and human disease

Mutations in SMCHD1 have been implicated in two human disorders, Facioscapulohumeral muscular dystrophy type 2 (FSHD2) and Bosma arrhinia microphthalmia (BAMS), associated with impaired regulation of the D4Z4 macrosatellite array ((Gordon et al., 2017; Lemmers et al., 2012; Shaw et al., 2017), for a review see (Jansz et al., 2017)). The D4Z4 array has a variable number of copies of the 3.3kb D4Z4 repeat unit, with each of them encoding for the TF DUX4 whose expression is limited to germ cells and epigenetically silenced in somatic cells.

FSHD patients inherit the allelic variant of the distal region of the D4Z4 array called 4qA, which encodes for a polyadenylation signal that stabilizes the *DUX4* transcript, and present variegated DUX4 expression due to loss of repressive chromatin modifications, such

as DNA hypomethylation (Jansz et al., 2017). In turn, DUX4 leads to activation of loci such as germline genes, endogenous retrotransposons and genes encoding for suppressors of differentiation in skeletal muscle which cause the onset of the pathology (Geng et al., 2012). There are two subtypes of FSHD categorised based on the genetic mutation that facilitates the loss of the heterochromatic landscape across the D4Z4 array. FSHD1 patients present a shortened D4Z4 array, while FSHD2 patients present, in most cases, a heterozygous mutation in *SMCHD1* (Jansz et al., 2017; Lemmers et al., 2019; Lemmers et al., 2012). Moreover, a *SMCHD1* mutation has a modifier effect in FSHD1 patients, causing a higher clinical severity than patients without such an alteration (Sacconi et al., 2013). FSHD2 mutations have been classified as loss of function and have been found across the full length of *SMCHD1*, although missense variants tend to be more often present in the ATPase and extended ATPase domains (Lemmers et al., 2019). Assessment of the *in vitro* ATPase activity of recombinant N-terminal mouse SmcHD1 constructs showed FSHD2 mutations cause a decrease in the catalytic activity of the protein (Gordon et al., 2017; Gurzau et al., 2018).

BAMS is a congenital disease characterized mainly by the absence of the nose and reduction in eye size. BAMS patients present hypomethylation of the D4Z4 array and are associated with missense mutations in *SMCHD1* which are exclusively localized in the ATPase and extended ATPase domains (Gordon et al., 2017; Shaw et al., 2017). Investigation of the effect of BAMS mutations on the *in vitro* ATPase activity of SmcHD1, as described above for FSHD2, showed these mutants have an increased ATP hydrolysis, suggesting they are gain of function mutations (Gordon et al., 2017; Gurzau et al., 2018). Moreover, injection in *Xenopus* embryos of mRNA encoding for mutant human *SMCHD1* caused craniofacial anomalies including microphthalmia, which interestingly could be induced also by overexpression of WT *SMCHD1*, hinting that BAMS mutations act by enhancing the protein's function (Gordon et al., 2017; Gurzau et al., 2018).

Binding of SMCHD1 to the D4Z4 array was shown by ChIP (Dion et al., 2019; Lemmers et al., 2012). Reminiscent of its action over the Xi and other autosomal targets, it was proposed the protein affects the DNA methylation status of the array. More specifically, it seems to play a role in its *de novo* methylation, as SmcHD1's removal in somatic cells did not cause hypomethylation, while the D4Z4 remethylation normally seen upon reprogramming of primary fibroblasts was impaired in cells coming from patients affected by FSHD2 or BAMS (Dion et al., 2019).

The uncovering of the structure of the ATPase and hinge domains of SMCHD1 has revealed the positioning of many of the disease-associated mutations in 3D space which is important for understanding better how they might influence aspects such as the structure, dimerization and ATPase activity of the protein (Chen et al., 2020; Pedersen et al., 2019). Notably, the exact pathway through which mutations in SMCHD1 lead to onset of either of these two disorders remains unknown. Studying it is challenging due to difficulties such as the lack of naturally occurring animal models (the D4Z4 array is primate-specific, (Clapp et al., 2007) and low availability of patient samples in the case of BAMS, which is a rare disease.

1.4 Aims of this study

While SmcHD1 has been proposed to be involved in several key processes associated with XCI, exactly how this protein mediates gene silencing, DNA methylation and chromosomal organization is currently unknown. Understanding its mechanism of action will be instrumental for gaining a more comprehensive understanding of not only XCI, which is a vital step in development of female mammals, but also of the pathogenesis of FSHD2 and BAMS through SmcHD1's role in regulating the repression of DUX4.

Most of the SmcHD1 studies discussed above have employed model systems containing or being equivalent to an SmcHD1 knockout which hasn't allowed for an understanding of the molecular mechanisms which enable SmcHD1 to interact with the underlying chromatin and perform these functions. Furthermore, these investigations have mainly relied on the analysis of averaged measurements taken from the behaviour of large populations of cells. Also, trials performed in the Brockdorff lab to perform SmcHD1 ChIP-seq using available antibodies have not been successful (Dr. Tatyana Nesterova, personal communication). Thus, in my work I have taken a different approach to investigate the mechanism of action of SmcHD1 by using live-imaging and single-particle tracking (SPT) to study the dynamics of SmcHD1. I aimed to answer the following questions:

- Is the dynamic behaviour of SmcHD1 on the Xi different to the rest of the nucleus?
- Does SmcHD1's dynamic behaviour over the Xi change as XCI progresses?
- How can SmcHD1's dynamic behaviour be reconciled with what we know about the recruitment and functions of the protein on the Xi?
- How do FSHD2- and BAMS-associated mutations alter the dynamic behaviour and functions of SmcHD1?

Chapter 2 – Materials and methods

2.1 Plasmids and cloning

2.1.1. Homology vectors cloning for CRISPR-Cas9 targeting

Homology arms of approximately 500-600bp flanking the insertion site were amplified by PCR from mESC genomic DNA using Fast-Start High Fidelity enzyme (Merk Life Science). The Halo- and eGFP-tags were amplified from previously generated vectors containing these sequences, using Fast-Start High Fidelity enzyme. The backbone plasmid was PCR amplified from previously generated vectors using Fast-Start High Fidelity enzyme. The fragments, homology arms and tag or homology arms only, were then ligated together with the backbone plasmid to generate the homology vector (HV) by Gibson Assembly ligation using Gibson Assembly Master Mix (NEB) according to the manufacturer's guidance. The ligated product (2µl) was transformed into XL10-Gold ultracompetent bacteria (Agilent) and DNA was isolated from single bacterial colonies using the Miniprep kit (Qiagen). The plasmid sequences were checked via Sanger sequencing (Source BioScience). For most of the homology vectors, the primers for Gibson Assembly were designed so that they introduced mutations to the PAM recognition sequence of the corresponding sgRNAs and/or mutations to create restriction enzyme (RE) sites required for screening by PCR followed by RE digestion. When introduction of such mutations was not possible through the Gibson Assembly primers, the plasmids obtained were subsequently mutagenized using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent). The primers used for Gibson Assembly and for mutagenesis are listed in Table 1.

Table 1: Primers used for Gibson Assembly and site-directed mutagenesis

Sequence 5'→3'	Purpose
CTGGGCCTCATGGGCCTTCC	Amplify plasmid backbone
ATGCGGCCTTGACGGCCTTCC	Amplify plasmid backbone
GGCCAAAAGCTGACGTACCAAACGGAGTGAGCAAGATGGCAGAAATCGGTAC	Amplify Halo tag; contains overhang for Gibson assembly to form SmcHD1 Halo HV
CATAGTCTTCAGAAATGCTCAGGTGGCGACCGGTAGGCCGGAAATCTCGAGC	Amplify Halo tag; contains overhang for Gibson assembly to form SmcHD1 Halo HV
CCTCCGCTGCCATGGTGGCGACCGGTAGGCCGGAAATCTCGAGC	Amplify Halo tag; contains overhang for Gibson assembly to form Halo SmcHD1 HV
CCTTCTCTTTTTTCTCCCTCAACGTGAGCAAGATGGCAGAAATCGGTAC	Amplify Halo tag; contains overhang for Gibson assembly to form Halo SmcHD1 HV
GAAGGAAGGCCGTCAAGGCCGCATGTGAGAGCCCTAGATTCAATCCTCTG	Amplify 5' homology arm for SmcHD1_E147A targeting; contains overhang for Gibson assembly to form SmcHD1_E147A HV
GATTAACGCTGCAAGAGCAAATGCTGTAAG	Amplify 5' homology arm for SmcHD1_E147A targeting
CATTGCTCTTGACGCGTTAATCGACAATTCATTGTCAGCTACTTCTCG	Amplify 3' homology arm for SmcHD1_E147A targeting; contains overhang for Gibson assembly to form SmcHD1_E147A HV
AAAGGAAGGCCCATGAGGCCAGGAGTTAGTTCCAGGACAGCCAAAGTG	Amplify 3' homology arm for SmcHD1_E147A targeting; contains overhang for Gibson assembly to form SmcHD1_E147A HV
GTAAGAGATTGAAAAACAGTACTTAACCTCTCCGACCCAAA	Mutagenize SmcHD1_E147A_HV to create RE digestion site for screening
TTTGGGTCGGAGAGGTTAAGTACTGTTTTTCAATCTCTTAC	Mutagenize SmcHD1_E147A_HV to create RE digestion site for screening
GAAGGAAGGCCGTCAAGGCCGCATCTGTGATGGATCAGTGGGCAGAAAC	Amplify 5' homology arm for SmcHD1_G137E targeting; contains overhang for Gibson assembly to form SmcHD1_G137E HV
CAAAGGATTCTGTTCTTCACTCGCATAATAC	Amplify 5' homology arm for SmcHD1_G137E targeting

Table 1 (continued)

Sequence 5'→3'	Purpose
CGAGTGAAGAACAGAATCCTTTGCGTAAGTAACCTGCTCCCGCTCGTTTTG	Amplify 3' homology arm for SmcHD1_G137E targeting; contains overhang for Gibson assembly to form SmcHD1_G137E_HV
AAAGGAAGGCCCATGAGGCCAGGCACAGGCAAGAGAGTCTGTGGAAAC	Amplify 3' homology arm for SmcHD1_G137E targeting; contains overhang for Gibson assembly to form SmcHD1_G137E_HV
GATCTGCAACTCCAGTCTTTCTAGAATTCGTCGATCGAGGCTAGAAC	Amplify ROSA26 splice acceptor site and beginning of eGFP sequence; contains overhang for Gibson assembly to form ROSA26_eGFP_CIZ1_HV
CCTCGCCCTTGCTCACCATTCTAGGAATTCGATACCGTCGATCCCCAC	Amplify ROSA26 splice acceptor site and beginning of eGFP sequence; contains overhang for Gibson assembly to form ROSA26_eGFP_CIZ1_HV
TAAGGCTCAGTCTGCTAATGCTTCCTGTGCCTTCTAGTTGCCAGCCATC	Amplify polyadenylation signal sequence and 3' homology arm for the ROSA26 locus; contains overhang for Gibson assembly to form ROSA26_eGFP_CIZ1_HV
CTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAC	Amplify polyadenylation signal sequence and 3' homology arm for the ROSA26 locus; contains overhang for Gibson assembly to form ROSA26_eGFP_CIZ1_HV
CTTTTTTTCTCCCTCAACGTGAGCAAGATGGCTTCTATGACCGGCGG	Amplify triple T7tag-linker-Halotag sequence; contains overhang for Gibson assembly to form T7HaloSmcHD1_HV
CGCTGCCATGGTGGCGACCGGTAGTCCGCCTCCGCTACCGGAAATCTC	Amplify triple T7tag-linker-Halotag sequence; contains overhang for Gibson assembly to form T7HaloSmcHD1_HV
CTACCGGTCGCCACCATGGCAGC	Amplify 5' homology arm, plasmid backbone and 3' homology arm from Halo_SmcHD1_HV
GAAGGAAGGCCGTCAAGGCCGCATCTTGTAGAGTAACCAACCCCCATACTC	Amplify 5' homology arm for SmcHD1_Y353C targeting; contains overhang for Gibson assembly to form SmcHD1_Y353C_HV

Table 1 (continued)

Sequence 5'→3'	Purpose
CTTTTGGTCCATGAATACAGTAATG	Amplify 5' homology arm for SmcHD1_Y353C targeting
CTGTATTCATGGACCAAAAGGAAATGAAATAAGTACAGCCAAAGCAATTG	Amplify 3' homology arm for SmcHD1_Y353C targeting; contains overhang for Gibson assembly to form SmcHD1_Y353C_HV
AAAGGAAGGCCCATGAGGCCAGGGACAAGCAAAACCTAGTTCTGTTG	Amplify 3' homology arm for SmcHD1_Y353C targeting; contains overhang for Gibson assembly to form SmcHD1_Y353C_HV
AGAAATTTCTTTTTGGTGGGAGATTATGTGAATTATAATGTCATGTTAGTAAAGAAT	Mutagenize SmcHD1_Y353C_HV to create RE digestion site for screening
ATTCTTTACTAACATGACATTATAATTCACATAATCTCCCACCAAAAAGAAATTTCT	Mutagenize SmcHD1_Y353C_HV to create RE digestion site for screening
TTCCCCGAAAAGTGCCACCTGACGTTCTAACACTGAAGTGTTGGATGATGTCGGA	Amplify 5' homology arm for Xist_ΔPID targeting; contains overhang for Gibson assembly to form Xist_ΔPID_HV; designed by Dr. Greta Pintacuda
AGATGGGCTATCTATACTCCTGACATCCAGTCATAATGTTAC	Amplify 5' homology arm for Xist_ΔPID targeting; contains overhang for Gibson assembly to form Xist_ΔPID_HV; designed by Dr. Greta Pintacuda
TGTCAGGAGTATAGATAGCCCATCTATACCCCCTCCATACTGAG	Amplify 3' homology arm for Xist_ΔPID targeting; contains overhang for Gibson assembly to form Xist_ΔPID_HV; designed by Dr. Greta Pintacuda
GCAGCGAGTCAGTGAGCGAGGAAGCAAAATGTTCCAGTGCAGAGGTTTT	Amplify 3' homology arm for Xist_ΔPID targeting; contains overhang for Gibson assembly to form Xist_ΔPID_HV; designed by Dr. Greta Pintacuda

2.1.2 sgRNA plasmids cloning for CRISPR-Cas9 targeting

Selection of the sgRNAs required for recognition and subsequent double-strand break generation at the target loci by CRISPR-Cas9 was done using the online tool CRISPOR (Concordet & Haeussler, 2018). sgRNAs with a high specificity score (above 80) and as little off-targets in exons as possible were selected. Annealed dsDNA oligos matching the chosen sgRNA were cloned into the PX459 V2.0 backbone (Addgene plasmid #62988) following the single-step digestion-ligation Zhang lab protocol (Ran et al., 2013). Two microliters of ligated product was transformed into XL10-Gold ultracompetent bacteria (Agilent) and DNA was extracted from bacterial colonies using the Miniprep kit (Qiagen). Sanger sequencing (Source BioScience) was used to confirm the sequence of the cloned plasmid. The sgRNA sequences used are listed in Table 2.

Table 2: Sequences of sgRNAs and associated PAMs used for CRISPR/Cas9

Target sequence	Sequence (5'→3'; PAM sequence in bold)	Note
SmcHD1_N-terminus	GAAAAAAGAGAAGGCCCTG CGG	Designed and cloned by Dr. Adam Cawte
SmcHD1_C-terminus	GTCTCTTTCAGAAGGCCAAA AGG	Designed and cloned by Dr. Adam Cawte
ROSA26	CGCCCATCTTCTAGAAAGACT TGG	Designed by Dr. Neil Blackledge and cloned by Dr. Tatyana Nesterova
LRIF1_exon1	GAGGAAAAGTCAGGCTCGCGCT TGG	Designed and cloned by Dr. Mafalda Almeida
LRIF1_exon4	TGTCTTCCCTGGTTGTACTAT TGG	Designed and cloned by Dr. Mafalda Almeida
SmcHD1_G137	TATGAGTATTATGCGAGTGA AGG	
SmcHD1_E147	AACTCTGCAAGAGCAAATGCT TGG	
SmcHD1_Y353	AATATTGTTGAAAGGTCCAAT TGG	
Xist_ΔPID	TTGAGGTTCTATACCAGTTC AGG	Designed by Dr. Greta Pintacuda

2.2 Cell culture

2.2.1 Mouse embryonic stem cell (mESC) growth

mESCs were grown at 37°C, 5% CO₂ in ES medium: Dulbecco's Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% Fetal Calf Serum (ThermoFisher),

0.1mM non-essential amino acids, 2mM L-glutamine, 50µM β-mercaptoethanol, 100 U/ml Penicillin/100 µg/ml Streptomycin (all from Life Technologies) and 1000U/mL Leukemia Inhibitory Factor (LIF) (made in-house by Dr. Tatyana Nesterova). The cells were grown on gelatinized plates, in feeder-dependent conditions. Feeders were generated from mitomycin-inactivated (Sigma Aldrich) SNLPs (STO mouse fibroblasts expressing Neomycin, Puromycin resistance and Lif genes). The main parental cell line for all subsequently derived cell lines is the interspecific (129/sv - Cast/Ei) iXist-ChrX XX mESC line derived by Dr. Tatyana Nesterova (Nesterova et al., 2019). Briefly, it contains a tetracycline-inducible promoter (TetOP) on the *M.m. domesticus* (129/sv) Xist allele and a reverse tetracycline-controlled transactivator (rTA) expressed from the *Tigre* locus. For Xist induction, cells were treated with 1µg/ml doxycycline (dox) (Sigma). When needed, MG132 treatment was done at 10µM for 6h, by addition of the drug to the media.

2.2.2 Generation of stable cell lines using CRISPR-Cas9-mediated homologous recombination

A day before transfection, 250 000 cells/well were seeded in 6-well plates on feeders. Cells were transfected using Lipofectamine 2000 (ThermoFisher) according to the manufacturer's protocol. Transfected cells were passaged 24h post-transfection onto gelatinized 90mm Petri dishes with feeders and, 48h post-transfection, puromycin (3-5µg/ml) selection was applied. Cells were kept under selection for 48h followed by exchanging for fresh ES media every day for a further 7-9 days until individual colonies were picked. The clones were then expanded in 96-well plates for screening. A summary of the cell lines generated is shown in Table 3.

Table 3: Summary of cell lines derived in this study and associated transfection and screening strategies

Cell line	Parental cell line	Transfected plasmids	Screening strategy
SmcHD1-Snap: SmcHD1 is tagged at the endogenous locus with a C-terminal SNAP-tag	iXist-ChrX line	SmcHD1_SNAP_HV (designed and cloned by Dr. Adam Cawte), PX459_SmcHD1_C-terminus	PCR
Halo-SmcHD1: SmcHD1 is tagged at the endogenous locus with an N-terminal Halo-tag	Xist-MS2+ROSA26-MCP-mNeonGreen (generated by Dr. Adam Cawte)	Halo_SmcHD1_HV, PX459_SmcHD1_N-terminus	PCR
SmcHD1-Halo: SmcHD1 is tagged at the endogenous locus with an C-terminal Halo-tag	Xist-MS2+ROSA26-MCP-mNeonGreen (generated by Dr. Adam Cawte)	SmcHD1_Halo_HV, PX459_SmcHD1_C-terminus	PCR
T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1: SmcHD1 is tagged at the endogenous locus with a T7 and Halo-tag; CIZ1 with an eGFP-tag at the N-terminus is expressed as a transgene from the ROSA26 locus	iXist-ChrX line	T7HaloSmcHD1_HV, PX459_SmcHD1_N-terminus, ROSA26_eGFP_CIZ1_HV, PX459_ROSAsgRNA	PCR
LRIF1_KO: A large part of the <i>LRIF1</i> gene is deleted causing a knock-out of LRIF1	T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1	PX459_LRIF1_exon1, PX459_LRIF1_exon4	PCR and WB
SmcHD1 ^{G137E} : SmcHD1 has a point mutation at the endogenous locus causing substitution of amino acid 137 glycine with glutamic acid	T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1	SmcHD1_G137E_HV, PX459_SmcHD1_G137	PCR followed by RE digestion

Table 3 (continued)

Cell line	Parental cell line	Transfected plasmids	Screening strategy
SmcHD1 ^{E147A} : SmcHD1 is has a point mutation at the endogenous locus causing substitution of amino acid 147 glutamic acid with alanine	T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1	SmcHD1_E147A_HV, PX459_SmcHD1_E147	PCR followed by RE digestion
SmcHD1 ^{Y353C} : SmcHD1 is has a point mutation at the endogenous locus causing substitution of amino acid 353 tyrosine with cysteine	T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1	SmcHD1_Y353C_HV, PX459_SmcHD1_Y353	PCR followed by RE digestion
Xist_ Δ PID: Xist has its B-repeat and part of C-repeat deleted	T7Halo-SmcHD1 +ROSA26_eGFP-CIZ1	Xist_ Δ PID_HV, PX459_Xist_ Δ PID	PCR and IF
H2BHalo: Expresses a randomly integrated transgene encoding for Halo-tagged histone H2B	iXist-ChrX line	pCAG_hH2B_Halo (gift from R. Klose lab)	Live imaging after labelling with a Halo-tag ligand

2.2.3 ML differentiation

For *in-vitro* differentiation, the cells were pre-plated for 30 - 40min in non-gelatinized plates to remove feeder cells, due to feeders' faster attachment relative to mESCs. The collected cells in suspension, which are primarily mESCs, were counted and plated at a density of 2000 cells/cm² in gelatinised plates. The differentiating cells were grown for up to 7 days. From the pre-plating step onwards, cells were kept in EC10 + dox medium: Dulbecco's

Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% Fetal Calf Serum (ThermoFisher), 0.1mM non-essential amino acids, 2mM L-glutamine, 50µM β-mercaptoethanol, 100 U/ml Penicillin/100 µg/ml Streptomycin (all from Life Technologies) and 1µg/ml doxycycline (Sigma).

2.3 gDNA extraction and colony screening

gDNA for the screening of the picked clones was extracted from a 96-well plate onto which the clones were split without feeders. Cells were lysed with lysis buffer (10mM NaCl, 10mM Tris pH 7.5, 10mM EDTA pH 8.0, 0.5% Sarcosyl) supplemented with proteinase K (1mg/ml) right before extraction. The plate sealed with a sticky film was left on a rotating platform at 55°C overnight. The DNA was then precipitated with ice-cold 100% EtOH at -20°C for at least 2h. The EtOH was removed by gently inverting the plate and the DNA was washed three times with ice-cold 70% EtOH. The DNA was then resuspended in 100µl nuclease-free water and used as template for PCR. To ensure identification of locus specific insertions, primer pairs used for screening were located outside of the inserted/mutated/deleted region and at least one primer was located outside of the homology arms. For the cell lines where a tag was being introduced, homozygous clones were shortlisted based on appearance of a larger PCR product due to presence of the tag and a loss of a WT-sized band. For the cell lines where a point mutation was generated, the PCR product was subsequently digested with the chosen RE and homozygous clones were shortlisted based on appearance of two smaller bands due to complete RE digestion and disappearance of a strong uncut WT-sized band. For the LRIF1_KO cell line, clones were shortlisted based on the presence of a small sized band which can only appear when a deletion took place between the sites targeted by the two introduced sgRNAs. A list of primers used for screening is shown in Table 4.

Shortlisted clones were then expanded and confirmed as being homozygous and having the correct sequence by PCR and Sanger sequencing. Furthermore, Western Blot was carried out to confirm the expression of the protein, checking for an upshift in size upon introduction of a tag, or the disappearance of the protein in the case of the LRIF1_KO. The retention of both X chromosomes in the selected clones was checked by PCR and RNA-FISH for Xist.

Table 4: List of primers used for screening of colonies from 96well plate

Primer	Sequence 5'→3'	Note
SmcHD1_N-terminus_tag_FW	GGTCACGGAAGGAGGCTTGG	Designed by Dr. Tatyana Nesterova
SmcHD1_N-terminus_tag_RV	CATCACCTCCCTGAACCTGAGGTAC	
SmcHD1_C-terminus_tag_FW	CTGCAGTTTTATCTCAGATTACACTGTCC	Designed by Dr. Adam Cawte
SmcHD1_C-terminus_tag_RV	GAATTTTCCTTACACAAAAGACAAAACATGTTTG	Designed by Dr. Adam Cawte
LRIF1_KO_FW	GCCGCCCTCTCAACTATCTC	Designed by Dr. Nick Brideau
LRIF1_KO_RV	GTTCTGGTGGTGTCACTGGA	Designed by Dr. Nick Brideau
SmcHD1_G137E_FW	GAGGCTGAGGCAGGAAGGTAAGTTTG	
SmcHD1_G137E_RV	CTCTGAAGTCCAGTGTGCAGGACTGC	
SmcHD1_E147A_FW	GTCCAAGGTCTGCCTGGGTTACAAAG	
SmcHD1_E147A_RV	CAATAGGCTCACTCATATCACTGGTG	
SmcHD1_Y353C_FW	CAGTAAGCTACCAAAGATGATTCCTG	
SmcHD1_Y353C_RV	GACATGACATTCTCTGCGTGTTTGG	
Xist_WT_FW	ACCATACACACACAAGTATCAACC	Designed by Dr. Tatyana Nesterova
Xist_WT_RV	AGCTTACGTACCTCCATCTTTAT	Designed by Dr. Tatyana Nesterova
Xist_Tet_RV	TGACCTCCATAGAAGACACCG	Designed by Dr. Tatyana Nesterova
Xist_ΔPID_FW	GTTTCGCTTGGTACTTCACG	
Xist_ΔPID_RV	GCAAGGGTAGTATTAGGACC	

2.4 Imaging methods

2.4.1 Single particle tracking

Cells were plated on gelatinised 35 mm μ -dishes containing a #1.5 polymer coverslip (Ibidi) two days before imaging in the case of differentiated cells or one day before imaging for undifferentiated cells. In the latter case, after plating, 1 μ g/ml doxycycline was added to the dish and the cells were imaged 24h later. Halo-PA-JF549 ligand (HHMI Janelia) (Grimm et al., 2016) was used for labelling at 100nM for 15min at 37 °C. Cells were then washed twice with PBS (Sigma) and incubated for 30min in Fluorobrite DMEM (ThermoFisher) supplemented with 10% Fetal Calf Serum (ThermoFisher) and 1 μ g/ml doxycycline. After three more washes with PBS, cells were imaged in Fluorobrite DMEM supplement as above and additionally with 30mM HEPES pH 7.4 to maintain the pH of the sample medium.

SPT experiments were carried out with a custom TIRF/HILO microscope, described previously (Huseyin & Klose, 2021), with an EMCCD camera (Andor), laser module (iChrome MLE MultiLaser engine, Toptica Photonics), 100x 1.4NA objective (Olympus) and translational module (ASI) which allows for adjustment of the angle of the excitation beam between epi- and HILO-illumination to achieve a high signal-to-noise ratio.

For fast tracking, which allows visualisation of both diffusing and bound molecules, movies of 4000 frames, with 15ms/frame, were acquired with continuous 561nm laser excitation at 22mW intensity and 405nm laser excitation at varied intensities to maintain a low density of fluorescent particles. Importantly, for the Xi-focused acquisitions, an image of the Xi region(s) was taken before and after the movie acquisition with 488nm laser excitation at 2mW and 20ms camera exposure. For whole nucleus datasets, at least 20 movies were taken, each containing about three cells, and at least three replicates were performed for each

timepoint studied. For Xi-focused datasets, 50 movies were taken, each containing at least one Xi domain, and at least four replicates were performed for each timepoint studied.

2.4.2 Single particle tracking analysis

Single particle tracking datasets were analysed predominantly in MATLAB (MathWorks) using Stormtracker software (Uphoff et al., 2013), unless otherwise specified. Localization of single particles in each frame of the acquired movies was achieved with a precision of 25 nm by fitting an elliptical Gaussian point spread function to the detected molecules. The detection of the molecules was done based on an intensity threshold fixed for all movies taken within the same acquisition session but adjusted as needed from session to session after manually checking the detection efficiency. Once localized, the molecules found within a radius of 766 nm in consecutive frames of the movie were linked to form tracks; single frame gaps were allowed, to take into account for spontaneous fluorophore blinking. The tracks which contained a minimum of four frames were then used to calculate their apparent diffusion coefficient D^* .

$$D^* = MSD/(4\Delta t) - \sigma_{loc}^2/\Delta t$$

where $\Delta t = 15.26$ msec, $\sigma_{loc} = 40$ nm and MSD is the mean-squared displacement:

$$MSD = 1/(N - 1) \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2$$

where $N \geq 5$ as only tracks with at least four frames are considered.

The log2 values corresponding to all the tracks found across all movies and all replicates for each timepoint analysed were displayed as a histogram using Igor Pro 9 (WaveMetrics); the

histograms were normalised to the number of tracks analysed in each timepoint and fitted with a double Gaussian distribution function.

$$f(x) = y_1 + A_1 e^{-\left(\frac{x-x_1}{w_1}\right)^2} + y_2 + A_2 e^{-\left(\frac{x-x_2}{w_2}\right)^2}$$

The tracks with a $\log_2(D^*)$ smaller than -3 were considered to be representing bound molecules and the bound population was calculated as a percentage of the number of these tracks from the total analysed tracks for that timepoint.

For the Xi-focused datasets, a filtering step was added to attribute the localized molecules found within the Xi territory prior to linking these localized molecules into tracks. To do this, firstly, the two images of the Xi region, taken before and after the movie acquisition, were merged using ImageJ/Fiji. If the region was found to have significantly moved during the acquisition, that particular field of view was discarded from further analysis. Then, a custom ImageJ/Fiji script designed by Dr. Aleksander Szczurek was used to manually select the Xi region, in each of the merged images, as well as a random region, of the same size as the Xi region but drawn somewhere else in the nucleus; the script outputs files containing the xy coordinates of these regions. Such files were input in another custom-made R script (also designed by Dr. Aleksander Szczurek) used to filter the localized single molecules, keeping only the molecules present in each of the selected regions. The filtered molecules were then merged into tracks which were analysed as described above.

2.4.3 Live cell imaging

Cells were plated in gelatinised μ -slide 8 wells or 35 mm μ -dishes with a polymer coverslip (Ibidi) at least four hours before imaging. Before imaging, cells were labelled with Halo-JF549 ligand (Promega) at 100nM or SNAP-JF549 ligand (gift from Luke Lavis, HHMI

Janelia) at 500nM for 30min at 37 °C. Cells were then incubated for 30min in growing media and subsequently washed twice with PBS (Sigma). Cells were imaged in Fluorobrite DMEM (ThermoFisher) supplemented with 10% Fetal Calf Serum (ThermoFisher) and, if needed, 1µg/ml doxycycline and/or 0.1µg/ml Hoechst (ThermoFisher). Live cell imaging was performed using an Olympus IX83 microscope equipped with a 63x 1.4NA objective, a 1200 x 1200 px² sCMOS camera (Photometrics 95B) and a humidified chamber with carbon dioxide atmosphere at 37°C. The microscope was operated via CellSens software. When needed, an additional magnifying 1.6x or 2x lens was used in front of the camera leading to a final pixel size of 114.4 nm or 91.5nm, respectively. z-stacks were captured every 0.5 µm, using 10-250ms exposure at 3-50% laser power depending on the channel. All acquired images were analysed and processed using ImageJ/Fiji.

2.4.4 Time course setup and analysis

For live imaging of the 7 day ML differentiation timecourse, all analysed cells were seeded for differentiation at the same time. Each day of the timecourse, cells were plated on µ-slide 8 wells with a polymer coverslip (Ibidi) four hours before imaging. The images were acquired with the Olympus IX83 microscope using the 2x lens in front of the camera. The analysis was done in ImageJ/Fiji using maximum-intensity projections of the acquired images. A binary mask of the Xi domains was created by thresholding the green channel images representing eGFP-CIZ1 signal. Each red channel image representing Halo-JF549-SmcHD1 signal was multiplied by its corresponding binary mask and then the Analyse Particles function was used to determine the mean SmcHD1 signal intensity in the each analysed Xi domain. The values for each day were plotted as a box plot using Igor Pro 9.

2.4.5 FRAP

Differentiating cells were plated on gelatinised 35 mm μ -dishes with a polymer coverslip (Ibidi) two days before imaging. Trolox 1mM (Sigma-Aldrich) was added to the media 24h before imaging to reduce phototoxicity as a result of prolonged imaging. Before imaging, cells were labelled as described in section 2.4.3 with Halo-JF549 ligand and imaged in Fluorobrite DMEM supplemented with 10% Fetal Calf Serum and 1 μ g/ml doxycycline. Live-cell timecourse imaging was performed using an Olympus SoRa spinning disc confocal equipped with a 63x 1.4NA objective, sCMOS camera (Photometrics), 405nm FRAP module and a humidified chamber with carbon dioxide atmosphere at 37°C. The microscope was operated using the CellSens software. The ROIs, either Xi or nucleoplasmic regions, were manually selected using the two-point circle measuring tool and photobleached with one bleach pulse of 40% laser intensity, following acquisition of two pre-bleached images with a delay of 1 minute between them. After photobleaching, images were taken every minute for a total of 20 minutes. Each image was taken with 100ms camera exposure with 10% 568nm laser intensity and consisted of a z-stack of 33 slices captured every 0.5 μ m. The acquisition was done on three different days for each cell line analysed.

2.4.6 FRAP analysis

Maximum intensity z-projections were used and the fluorescence loss due to photobleaching across the timecourse was corrected for using the bleach correction plugin in ImageJ/Fiji, fitting to an exponential decay. The fluorescence intensity measured within the bleached region was normalised to the background intensity, measured outside of the cell, and the fluorescence recovery was then represented as a percentage relative to the fluorescence intensity pre-bleaching. Curves showing the averaged recovery of all cells analysed for each

condition as a function of time were plotted using Igor Pro 9. The curves were then fitted with a single exponential function in Igor Pro 9 in order to determine the immobile population and lifetime.

$$f(t) = y_0 + Ae^{\frac{-(t-t_0)}{\tau}}$$

where t is time, y_0 is the plateau of the curve, A and τ are coefficients and t_0 is a constant and equals 0.

2.4.7 Immunofluorescence

Cells were plated on gelatinised 35 mm μ -dishes or coverslips two days before fixation. After two washes with PBS, cells were fixed in 2% paraformaldehyde for 15min at RT and permeabilized in 0.4% Triton X-100 for 5min at RT. Cells were briefly washed with PBS, blocked for three periods of 2min with 0.2% fish gelatine in PBS and incubated with primary antibody (anti-SmcHD1 1:1000 (made in-house), anti-H3K27me3 1:1000 (Active Motif, # 61017) diluted in fish gelatine solution with 5% normal goat serum for 2h at RT in a humid chamber. Samples were washed three times with fish gelatine solution and incubated for 1h at RT with secondary antibody (Alexa Fluor 568 goat anti-mouse IgG, Alexa Fluor 488 goat anti-rabbit IgG) 1:500 diluted in fish gelatine solution. Following two washes with fish gelatine solution and one wash with PBS, Vectashield with DAPI was added and the μ -dish lid was locked with parafilm or the coverslips were sealed to slides with nail varnish. Cells were analysed on the Olympus IX83 microscope described in section 2.4.3 or an inverted fluorescence Axio Observer Z.1 microscope using a PlanApo x63/1.4 NA oil-immersion objective.

2.4.8 H3K27me3 domain size analysis

Z-stack images taken with the Olympus IX83 microscope described in section 2.4.3 were analysed using the 3D Objects Counter function in Fiji/ImageJ. For each image, a threshold was set in order to highlight the H3K27me3 focus in each cell and the volume of each focus was chosen as an output. The object map was cross-checked against the original image to make sure the foci have been correctly identified. For each cell line, the volume values were plotted as a violin plot using Igor Pro 9.

2.5.1 Chromatin-associated RNA extraction and sequencing (chrRNA-seq)

mESCs growing without feeders or ML-differentiated cells were pelleted from confluent 90mm dishes, washed once with PBS and snap-frozen for storage at -80°C. The pellets were resuspended in HLBN buffer (0.01M NaCl, 2.5mM MgCl₂, 10mM Tris pH 7.5, 0.1% NP-40). Then the cells were underlaid with HLBN buffer supplemented with 24% sucrose and nuclei were isolated through centrifugation (1000g 5min at 4°C). Nuclei were resuspended in NUN1 (75mM NaCl, 20mM Tris pH 7.9, 0.5mM EDTA, 50% glycerol, 0.1mM DTT) and then lysed with NUN2 (0.3M NaCl, 20mM HEPES pH 7.6, 7.5mM MgCl₂, 0.2mM EDTA, 1M urea, 1% NP-40, 0.1mM DTT) by incubating 15min on ice and occasionally vortexing. The chromatin was isolated by centrifugation (15000rpm 10min at 4°C) and resuspended in HSB buffer (0.5M NaCl, 10mM Tris pH 7.5, 10mM MgCl₂). The samples were DNase (LifeTech) treated at 37°C, shaking (1400rpm) until the chromatin pellet dissolved. Further, the samples were treated with 0.2% SDS and 0.4mg/ml proteinase K for 30min at 37°C, shaking (1400rpm). To obtain the chromatin-associated RNA, two rounds of standard TRIzol- (Invitrogen) chloroform extraction with overnight isopropanol precipitation were performed. Between the two rounds, a second DNase treatment was done for 30min at 37°C,

shaking (1400rpm). After the second round, the RNA pellets were resuspended in RNase-free water and their concentration was measured by NanoDrop (ThermoFisher). 500ng of RNA was used for library preparation with the Illumina TruSeq stranded total RNA kit (RS-122-2301). The pooled libraries were sequenced by 2 x 81 paired-end sequencing using Illumina NextSeq500 (FC-404-2002).

2.5.2 ChrRNA-seq analysis (Note all chrRNA-seq analysis was done by Dr. Guifeng Wei)

The analysis of the chrRNA-seq datasets was done as previously described (Nesterova et al., 2019). Briefly, raw fastq files were mapped using bowtie2 (v2.3.2, (Langmead & Salzberg, 2012)) to rRNA and the rRNA-mapped reads were discarded. Subsequently, the unmapped reads were aligned to an N-masked mm10 genome with STAR (v2.4.2a or v2.7.9a, (Dobin et al., 2013)) with parameters “*--twopassMode Basic --outSAMstrandField intronMotif --outSAMattributes All --outFilterMultimapNmax 1 --outFilterMismatchNoverReadLmax 0.06 --alignEndsType EndToEnd*”. Unique alignments were then split into the two alleles, Cast and 129S, using SNPsplit (v0.4.0dev, (Krueger & Andrews, 2016)). featureCounts program, with parameters “*-t transcript -g gene_id -s 2*”, (Liao et al., 2014)) was used to count the allelic and unsplit reads that overlap with genes. Samtools (Li et al., 2009) was used to filter and sort the alignments. Only genes with at least 10 SNP-covering reads across samples for comparisons were used to calculate the allelic ratio $X_i/(X_i+X_a)$ (X_i represents the inactive allele which here is the 129 one, while X_a represents the active allele which here is the Cast one). The genes were split into three categories according to their dependency on SmcHD1 for silencing as previously defined in (Gdula et al., 2019). Boxplots were generated using R package ggplot2. The designed mutations (i.e. G137E, E147A, Y353C) for SmcHD1

and deletions for knockouts are confirmed by the chrRNA-seq reads and are visualised using IGV.

2.6 Whole cell lysates, RIPA extracts, nuclear extracts and western blot

Whole cell lysates were obtained from mESCs growing without feeders, by washing the cells with PBS and then lysing them with SMASH buffer (0.125M Tris pH 6.8, 11.6% glycerol, 4% SDS, 0.002% bromophenol blue, 10% β -mercaptoethanol). The lysates were snap-frozen and stored at -80°C until use.

Whole cell extracts were obtained from mESCs growing without feeders. Cells were washed once with PBS and then the extraction was done using RIPA buffer (Cell Signalling 9806) supplemented with 1mM PMSF, 1x complete protease inhibitors (Roche), 0.5 μ l/ml benzonase (Merk E8263-5KU) and 2mM MgCl₂ on ice for 30-60 minutes. After snap-freezing and thawing, the samples were centrifuged and the supernatant was kept as the soluble whole cell extract which was snap-frozen and stored at -80°C until use. The concentration of the extract was determined using the Bio-Rad Bradford assay according to the manufacturer's protocol.

For nuclear extracts, mESCs growing without feeders were pelleted, washed with PBS, snap-frozen on dry ice and stored at -80°C. The pellets were resuspended in hypotonic buffer A (10mM HEPES pH 7.9, 1.5mM MgCl₂, 10mM KCl) supplemented with 0.5mM DTT, 0.5mM PMSF and 1x complete protease inhibitors (Roche) and incubated on ice for 10min. Cells recovered by centrifugation, were lysed in buffer A with 0.1% NP-40 supplemented with 5mM DTT, 0.5mM PMSF and 1x complete protease inhibitors and incubated on ice for 10min. The nuclei were then isolated by centrifugation, washed once with PBS

supplemented with 1x complete protease inhibitors, and resuspended in high salt buffer C (5mM HEPES pH 7.9, 26% glycerol, 1.5mM MgCl₂, 0.2mM EDTA, 350mM NaCl) supplemented with 0.5mM DTT and 1x complete protease inhibitors. The extraction was done on ice for 1h and the nuclei were then pelleted with the supernatant kept as the nuclear extract which was snap-frozen and stored at -80°C until use. The concentration of the extract was determined using the Bio-Rad Bradford assay according to the manufacturer's protocol. To analyse the chromatin insoluble fraction, the last nuclear pellet obtained was resuspended in SMASH buffer, snap-frozen and stored at -80°C until use.

The extracts obtained through one of the methods above were separated on a home-made SDS-PAGE separating gel with a stacking gel or a commercial NuPAGE gel (ThermoFisher). The proteins were transferred onto a nitrocellulose membrane (Bio-Rad) using the Bio-Rad Trans-Blot Turbo Transfer system, with High MW or Mixed MW preprogrammed protocols depending on the protein size. Membranes were blocked for 1h at RT in PBST (PBS with 0.1% Tween 20) with 5% milk and incubated with primary antibodies (summarised in Table 5) diluted in blocking solution, overnight shaking at 4°C. Following washes with PBST, membranes were incubated for 1h at RT with secondary antibodies conjugated to horseradish peroxidase (Amersham) diluted 1:2000 in blocking solution. Membranes were washed with PBST and the bands were revealed using enhanced chemiluminescence (ECL) solution (ThermoFisher) and developed on chemiluminescent sensitive photographic film (Amersham).

Table 5: Antibodies used for Western Blot

Antibody	Host	Source	Dilution
SmcHD1	rabbit	Made in house	1:1000
HDAC3	rabbit	Abcam, ab32369	1:1000
METTL3	rabbit	Abcam, ab195352	1:10000
CIZ1	rabbit	Gift from D. Coverley	1:1000
LRIF1	rabbit	Proteintech, Cat. 26115-1-AP	1:5000
H2AK119ub1	rabbit	Cell Signalling, 8240S	1:2000
TBP	mouse	Abcam 51841	1:2000

Chapter 3: Investigating dynamic behaviour of SmcHD1 on the inactive X chromosome using single-particle tracking

3.1 Introduction

SPT is a powerful method that allows one to study the dynamic behaviour of a protein by live-cell imaging. Diffusive behaviour is measured by following single fluorescently labelled molecules over time, with data from multiple molecules being used to define sub-populations, for example molecules with diffusive vs sub-diffusive dynamics. Typically, proteins of interest are labelled by attaching a tag which can be either a photoactivable fluorescent protein or a self-labelling tag, such as Halo- or Snap-tag, that can be paired with photoactivable ligands. Photoactivable fluorophores need to be illuminated first with a laser to become photoactivated and, subsequently, with an excitation laser. This gives the user control over the number of fluorescent molecules at any point during acquisition. It is essential to avoid high densities of particles to be able to accurately locate and assign each particle in each frame to their respective track at the analysis stage (Uphoff et al., 2014). Importantly, to have a high enough signal-to-noise ratio to visualize single particles, imaging is usually performed using highly inclined and laminated optical sheet (HILO) illumination, where the cells are illuminated by a thin laser beam at an angle (Tokunaga et al., 2008).

In recent years the application of SPT has been key in understanding the mechanism of action of chromatin-binding proteins. For example, in a study by Szczurek et al. (2024), SPT was used to understand how the Polycomb system facilitates gene repression. Several factors involved in the initiation and early elongation stages of transcription were Halo-tagged and changes in their dynamic behaviour, particularly in the proportion of their bound molecules and for how long they are stably residing on chromatin, were determined in the presence and absence of PRC1 (Szczurek et al., 2024). This analysis revealed that the Polycomb system prevents binding of components of the general TF TFIID which plays a role in the early stages of transcription initiation (Szczurek et al., 2024).

SPT has also been used in the context of XCI, where an Xi-focused study was performed to resolve a long-standing question regarding the relationship between RNAPII and the Xi territory (Collombet et al., 2023). Immunofluorescence analysis has shown that the transcription machinery is excluded from the Xi region immediately after the onset of XCI (Chaumeil et al., 2006; Fraser & Bickmore, 2007) and it has been suggested that this could be due to the Xi being a physically a separated compartment (Cerase et al., 2019). This study shows the reason behind the apparent exclusion is actually a greatly decreased bound population of the polymerase over the Xi compared to the nucleoplasm, rather than the presence of any barrier against RNAPII diffusion into the domain (Collombet et al., 2023). The authors propose that Xist, together with the other XCI players, act to impede RNAPII from stably binding across the Xi territory to prevent initiation of active transcription, although the exact mechanism remains to be determined (Collombet et al., 2023).

In this chapter I describe an experimental setup which allows cellular imaging and SPT of endogenous SmcHD1, and, crucially, specific analysis of SmcHD1 molecules from within the Xi domain. I apply this model to study changes in the dynamic behaviour of SmcHD1 during XCI. Subsequent chapters 4 and 5 will follow how the WT dynamics are

affected by the introduction of various mutations in the system and the effects of these mutations on the protein's functionality.

3.2 Generation of a system for SmcHD1 single-particle tracking

I elected to use as a model system XX female mESCs which have inducible Xist RNA expression, in order to have control over the onset of XCI. A suitable cell line, iXist-ChrX (Figure 3.1 A), has been previously generated in the lab by Dr. Tatyana Nesterova using CRISPR/Cas9-mediated genome engineering. More specifically, on one of the Xist alleles, the endogenous promoter has been replaced with a tetracycline-dependent promoter whose activation is controlled by a transactivator rTA that has been targeted to be constitutively expressed under a pCAG promoter from the *Tigre* locus (Gossen et al., 1995). Addition of doxycycline to the growth media leads to synchronised Xist expression in about 95% of cells (Nesterova et al., 2019). Also, iXist-ChrX is an interspecific *M.m. domesticus* (129S) x *M.m. castaneus* (Cast) cell line which allows for assignment of sequencing reads to their strain of origin based on the presence of SNPs. This is key for performing allele-specific analysis of X-linked transcription, which I use in this study (see Chapters 4 and 5). However, it is important to note that an unintentional chromosomal rearrangement during derivation of this cell line limits allele-specific analysis to a 103Mb region proximal to Xist (Nesterova et al., 2019).

Using the iXist-ChrX line as a parental cell line, I used CRISPR/Cas9-mediated homologous recombination to homozygously target the C-terminus of endogenous *SmcHD1* with a Snap-tag (Figure 3.1 A). I chose to tag the C- rather than the N-terminus in order to avoid affecting the functionality of the N-terminal GHKL ATPase domain which has been

shown to be required for enrichment of the protein over the Xi domain (Brideau et al., 2015; Gurzau et al., 2021). I opted to use the Snap-tag since it is a self-labelling protein tag that allows visualization of the protein of interest with a series of fluorescently-labelled dyes, including photoactivable dyes which are key for SPT (Keppler et al., 2003; Uphoff et al., 2014). The plasmids required for this targeting were already available in the lab, having been previously cloned by Dr. Adam Cawte. The obtained clones were originally screened by PCR (data not shown), and then expression of the tagged protein was checked for the shortlisted clones by western blot with an anti-SmcHD1 antibody (Figure 3.1 B). WT SmcHD1 has a molecular weight of 226kDa and as expected, introduction of the Snap-tag causes a shift of 20kDa in all three clones. Notably, the level of the tagged protein is markedly lower than of the WT protein. During the generation of this cell line, a colleague in the lab, Dr. Tatyana Nesterova, attempted a similar targeting of the C-terminus of SmcHD1 but introduced a T7-tag instead, which is much smaller than the Snap-tag, around 4kDa. She observed the same decrease in protein levels as reported here for SmcHD1-Snap suggesting that the reduction is a consequence of targeting the C-terminus of the protein regardless of the tag (personal communication). However, chrRNA-seq carried out using the C-terminal T7-tagged SmcHD1 showed normal X-linked gene silencing, indicating that the reduced protein levels do not affect its functionality (Figure 3.1 F). Accordingly, I proceeded to use the SmcHD1-Snap cell line for live-cell imaging.

Surprisingly, upon Snap-tag labelling in undifferentiated mESCs, SmcHD1 was observed to be enriched over Hoechst-dense pericentric heterochromatin regions (Figure 3.1 C). This was unexpected given that previous visualisation of SmcHD1 in ESCs by immunostaining showed an even distribution of the protein across the nucleus with pericentric heterochromatin staining detectable in only in a very small proportion (about 14%) of cells, (Brideau et al., 2015; Jansz, Nesterova, et al., 2018). Immunostaining carried

out in the same cell line used for live imaging, also in undifferentiated mESCs, showed indeed the pattern previously observed (Figure 3.1 D). The pathway which enables localization of SmcHD1 to the pericentric regions is further explored in Chapter 4. However, the presence of these unanticipated foci across the nucleus makes it difficult to define the position of the Xi (Figure 3.1 E), suggesting I would need an independent marker for the location of the Xi chromosome in live-cell imaging. Moreover, initial trials for using the SmcHD1-Snap cell line in HILO illumination for SPT showed that the signal may not be bright enough to achieve the required signal-to-noise ratio (data not shown). Taking the above into account, I decided to establish a modified cell line model which would mitigate the limitations of the SmcHD1-Snap mESC line.

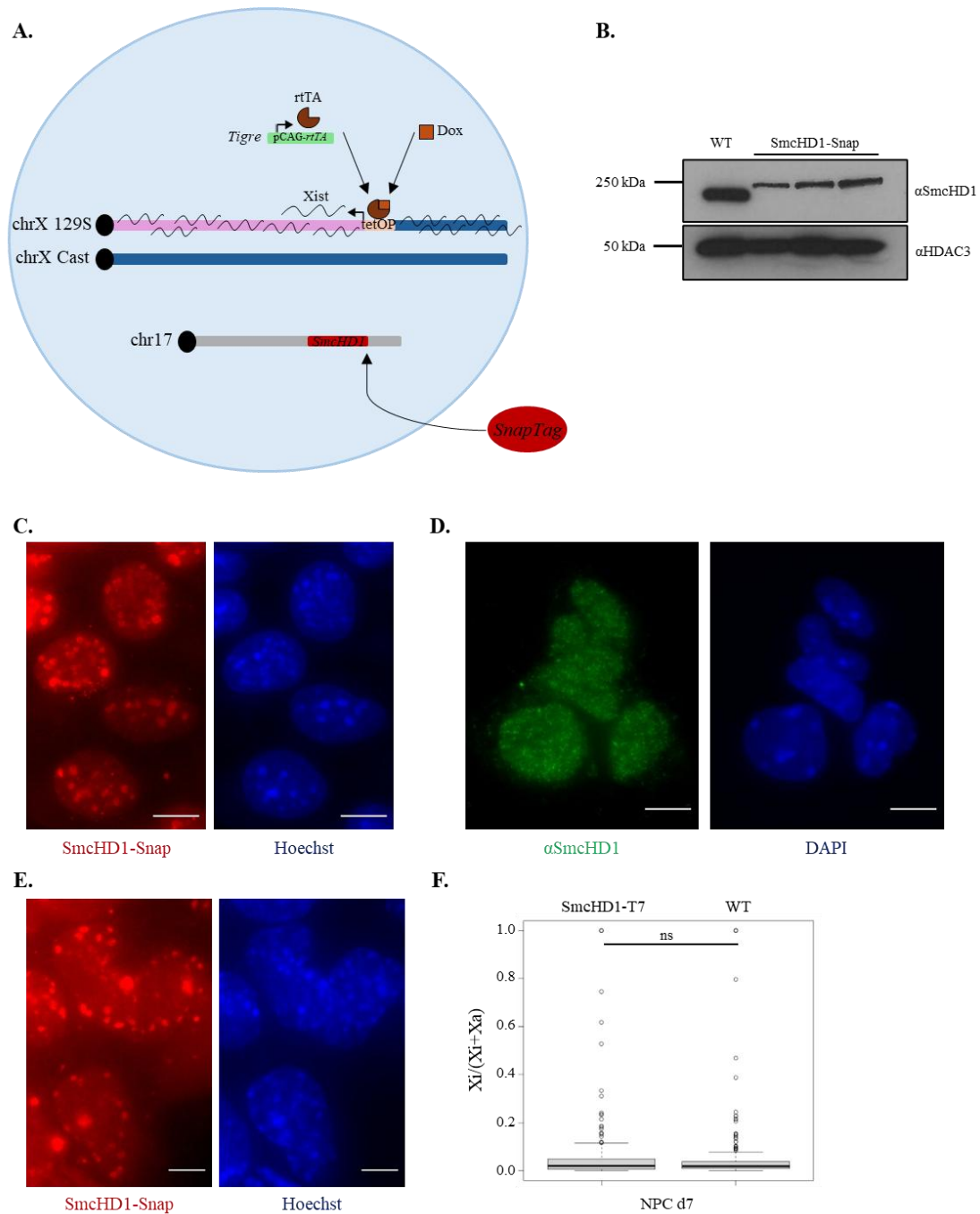


Figure 3.1 – Generation of a mESC line with endogenously Snap-tagged SmcHD1:

A. Schematic showing the interspecific iXist-ChrX XX mESC line used as a parental cell line throughout this project and the introduction of a Snap-tag at the C-terminus of endogenous SmcHD1. **B.** Western Blot from nuclear extract showing the expression level of SmcHD1-Snap in three different clones compared to WT SmcHD1; HDAC3 acts as a loading control. **C.** Representative live images of Snap-tag labelling and Hoechst staining in SmcHD1-Snap mESC line, in absence of doxycycline addition. Images are maximum-intensity projections, scale bar corresponds to 10μm. **D.** Immunostaining for SmcHD1 in SmcHD1-Snap mESC line, in absence of doxycycline addition. Images are maximum-intensity projections, scale bar corresponds to 10μm.

Figure 3.1 (continued) – Generation of a mESC line with endogenously Snap-tagged SmcHD1:

E. Representative live images of Snap-tag labelling and Hoechst staining in SmcHD1-Snap mESC line, taken at day 7 of ML differentiation. Images are maximum-intensity projections, scale bar corresponds to 5µm. **F.** Box plot showing allelic chrRNA-seq analysis of transcription of X-linked genes in WT and SmcHD1 C-terminal T7-tagged mESCs which were NPC differentiated for 7 days. The WT sample represents an average of two replicates, the SmcHD1-T7 sample represents one replicate. Statistical analysis was done using the paired t-test.

3.3 Generation of an improved system which allows simultaneous visualisation of SmcHD1 and Xi domain in live imaging

For the modified model system I used as a parental line the same interspecific iXist-ChrX XX mESC line described in section 3.2 (Nesterova et al., 2019), but switched the Snap-tag for a Halo-tag since it has higher labelling efficiency (Figure 3.2 A). Previous comparison between the two tags in SPT has suggested the Halo-tag is more photostable and presents less unspecific binding (Latty et al., 2015; Presman et al., 2017). I also introduced a triple T7-tag in case it would be needed for further assays (I didn't make use of this in this study). Using CRISPR/Cas9-mediated targeting, I introduced the T7-Halo-tag in the endogenous *SmcHD1* locus at either the N- or the C-terminus in order to compare directly the effect of the introduced tag on the protein's expression level. For the clones having the tag at the C-terminus, I observed the same reduction in expression previously seen with the Snap- and T7-tags (data not shown). Conversely, having the tag on the N-terminus didn't decrease SmcHD1 levels (Figure 3.2 B).

Halo-tag labelling in undifferentiated mESCs showed SmcHD1 association with pericentric heterochromatin as observed in the SmcHD1-Snap line (Figure 3.2 D).

Therefore, in order to discriminate Xi/Xist RNA I used a cell line having Xist tagged with MS2v5 loops and fluorescently tagged MS2 coat protein expressed from the *Rosa26* locus. Such a system has been developed for visualisation of RNAs, including Xist, in live-cell imaging in mammalian cells (Bertrand et al., 1998; Ng et al., 2011). However, in this case, the tagging resulted in ~4-fold reduction in Xist levels and subsequent impaired accumulation of SmcHD1 over the Xi domain and X-linked gene silencing (data not shown). I then used an alternative strategy, expressing a construct encoding CIZ1 eGFP-tagged at its N-terminus in the *Rosa26* locus. As described in Section 1.2.2.1, CIZ1 binds Xist to aid its localisation over the Xi chromosome and it can, thus be used as a proxy for Xist RNA sites. eGFP-CIZ1 expression was confirmed by western blot (Figure 3.2 C). Importantly, CIZ1 is enriched over the Xi coincident with the onset of Xist expression and continuously thereafter (Ridings-Figueroa et al., 2017). These characteristics together with the introduced eGFP-tag allow for visualisation of Xi domain in live-cell imaging, even in undifferentiated mESCs (Figure 3.2 D). From two clones which had integrated both the tag on SmcHD1 and the CIZ1 transgene, I selected one, marked with an asterisk in Figure 3.2 B-C, for all further experiments based on the proteins' expression level and quality of signal in live imaging.

Previous studies analysing SmcHD1 recruitment to the Xi using fixed samples and IF suggest SmcHD1 only enriches on the Xi in differentiated cells (Bowness et al., 2022). Visual inspection of the SmcHD1 signal underneath the CIZ1-marked Xi domain and checking the intensity profile of the SmcHD1 and CIZ1 signals across a line drawn through the Xi domain, an example of which can be seen in Figure 3.2 E, confirm that SmcHD1 is not enriched over the Xi chromosome in the absence of cellular differentiation, as previously shown (Bowness et al., 2022). Now with the development of a more sensitive live-cell imaging approach, the amount and timing of SmcHD1 enrichment can be more accurately determined. In this respect, I have carried out a live imaging timecourse over 7 days of ML

differentiation, a differentiation protocol which involves growing mESCs as a monolayer by plating them at a low density and removing LIF. I then quantified the intensity of the SmcHD1 signal beneath the CIZ1-signal across many individual cells (Figure 3.2 F-G). This analysis showed that SmcHD1 accumulates gradually over the Xi, ultimately forming a domain at which the protein is visibly enriched compared to the rest of the nucleus (Figure 3.2 F-H).

Having setup a system which allows visualisation of both SmcHD1 and the Xi domain in live imaging, I went ahead to study the dynamic behaviour of SmcHD1. Throughout this study I will be comparing two cellular states: day 0 that represents undifferentiated mESCs, which are not undergoing XCI, but which have been treated for 24h with doxycycline to promote Xist expression and, thus, allow accumulation of CIZ1 and visualisation of the future Xi's location; and day 7 that represents mESCs differentiated for 7 days, which are undergoing XCI. Importantly, at day 7, SmcHD1 is strongly enriched over the Xi domain, X-linked gene silencing is almost fully completed and any SmcHD1-dependent defective silencing can be distinguished, as seen, for example, in the case of an SmcHD1 knock-out (Bowness et al., 2022). Therefore, day 7 is a timepoint that will allow the pairing of live-cell imaging data with data assessing the functionality of various SmcHD1 mutants.

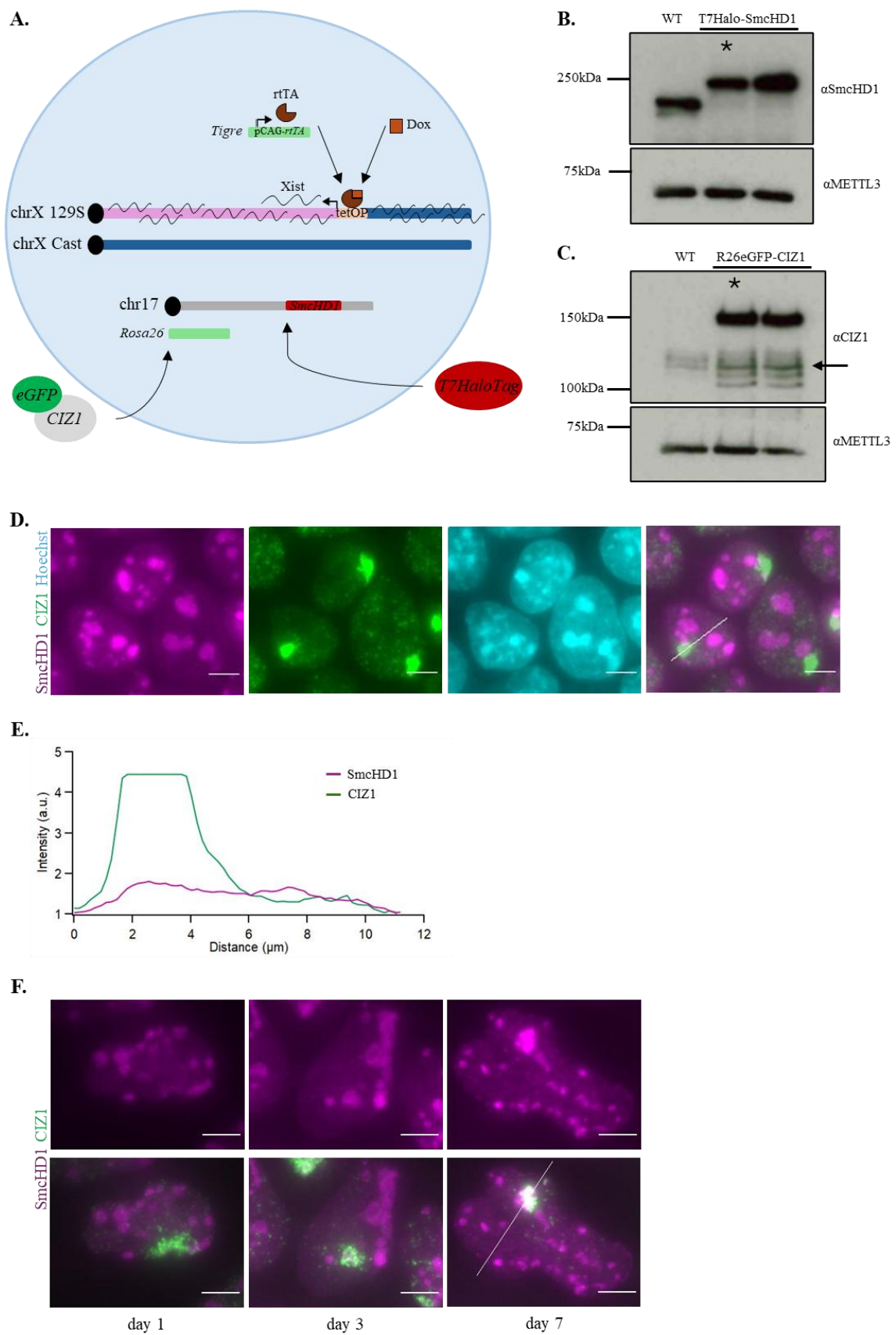


Figure 3.2 – Generation a mESC line with endogenously Halo-tagged SmcHD1 and eGFP-CIZ1 transgene expressed from the *Rosa26* locus

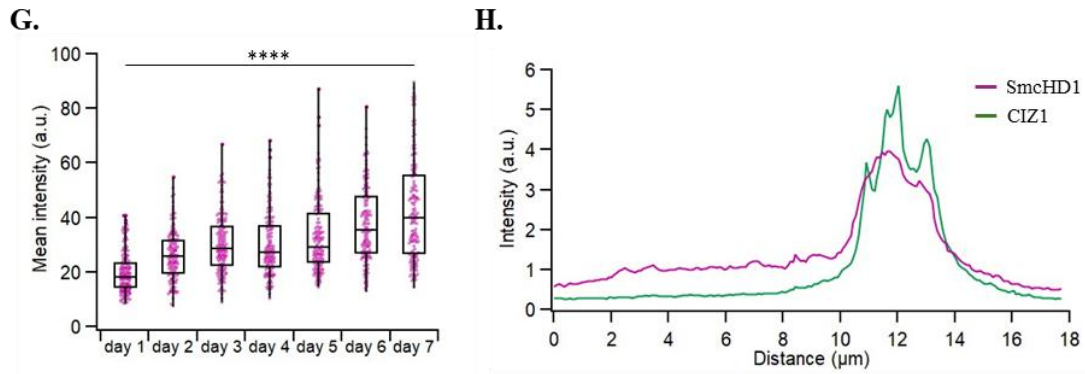


Figure 3.2 (continued) – Generation a mESC line with endogenously Halo-tagged SmcHD1 and eGFP-CIZ1 transgene expressed from the *Rosa26* locus:

A. Schematic showing the interspecific iXist-ChrX XX mESC line used as a parental cell line throughout this project with the introduction of a T7-Halo-tag at the N-terminus of endogenous SmcHD1 and the transgene encoding for eGFP-tagged CIZ1 into the *Rosa26* locus. **B.** Western Blot from nuclear extract showing the expression of T7Halo-tagged SmcHD1 in two different clones compared to WT SmcHD1; METTL3 acts as a loading control. **C.** Western Blot from the same nuclear extract as in B. showing the expression of WT CIZ1 (the specific band is marked by the arrow) and eGFP-tagged CIZ1 in the two different clones; METTL3 acts as a loading control. The * shows the clone that was used for further experiments and cell lines generations throughout this project. **D.** Representative live images of Halo-tag labelling, eGFP-CIZ1 accumulation over the Xi chromosome, Hoechst staining in T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line, taken after 24h doxycycline addition to induce Xist expression. The fourth image represents a merge between the Halo and eGFP signals and the line across the cell shows the position of the intensity profile measured in D. Images are maximum-intensity projections, scale bar corresponds to 5μm. **E.** Intensity profiles for SmcHD1 (magenta) and CIZ1 (green) across the line shown in D. **F.** Representative live images showing gradual accumulation of SmcHD1 over the Xi domain marked by eGFP-CIZ1 from day 1 to day 3 to day 7 of ML differentiation. Images are maximum-intensity projections, scale bar corresponds to 5μm. **G.** Quantification of SmcHD1 signal underneath the CIZ1 domain, at each day, during a 7 days live imaging ML differentiation timescourse. Each dot represents a cell. n = 136-159. Statistical analysis was done using the unpaired t-test. **H.** Intensity profiles for SmcHD1 (magenta) and CIZ1 (green) across the line shown in F.

3.4 Single-particle tracking to define SmcHD1 dynamic behaviour across the cell nucleus

In a first series of experiments, I determined SmcHD1 dynamics across the entire nucleus in undifferentiated mESCs and in differentiated cells. I took advantage of the introduced Halo-tag to label SmcHD1 with a photoactivable ligand, Halo-PA-JF549, and investigated its dynamic behaviour at day 0 (undifferentiated mESCs treated with doxycycline for 24h) and day 7 (mESCs differentiated for 7 days) by tracking single molecules of the protein, aided by the use of HILO microscopy with high temporal resolution (15ms frames). This was performed as described in Section 2.4.1, using a setup and size of datasets (at least 20 fields of view, containing around three cells each were analysed per replicate) as used in previous studies analysing the dynamic behaviour of chromatin interacting proteins (Huseyin & Klose, 2021; Szczurek et al., 2024). Looking at the histograms of the diffusion coefficients for the obtained tracks, two populations emerge in both conditions: an immobile population with a Log_2 diffusion coefficient peak around $-4 \mu\text{m}^2/\text{s}$, which represents SmcHD1 bound to chromatin and a larger, fast population with a peak around $-1 \mu\text{m}^2/\text{s}$, which corresponds to freely diffusing SmcHD1 (Figure 3.3 A-B). Such diffusion coefficient values and distributions are similar to those observed for cohesin subunit Scc1 and its loader Scc2 which are factors involved in chromosome organisation (Rhodes et al., 2017). Altogether, this confirms the suitability of the T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line for use in SPT studies.

A comparison of the histograms for undifferentiated and differentiated cells shows no striking changes between the two cellular states (Figure 3.3 C). To quantify the amount of bound molecules, I calculated the percentage of the number of tracks whose Log_2 diffusion

coefficient was smaller than -3 from the total number of tracks (see Section 2.4.2). -3 was chosen as a boundary value as it was seen, across the various timepoints and cell lines analysed in this study, to sit close to the valley between the peaks of the two populations. Thus, it can be determined that the bound population of SmcHD1 remains unchanged during differentiation, at about 42% (Figure 3.3 D). However, it is very important to point out that this analysis takes into account SmcHD1 molecules coming from all across the nucleus. This does not allow us to know whether any differences in SmcHD1 dynamic behaviour are due to its role in XCI or just general changes in its interaction with chromatin during differentiation. Thus, I went on to perform Xi-specific SPT analysis.

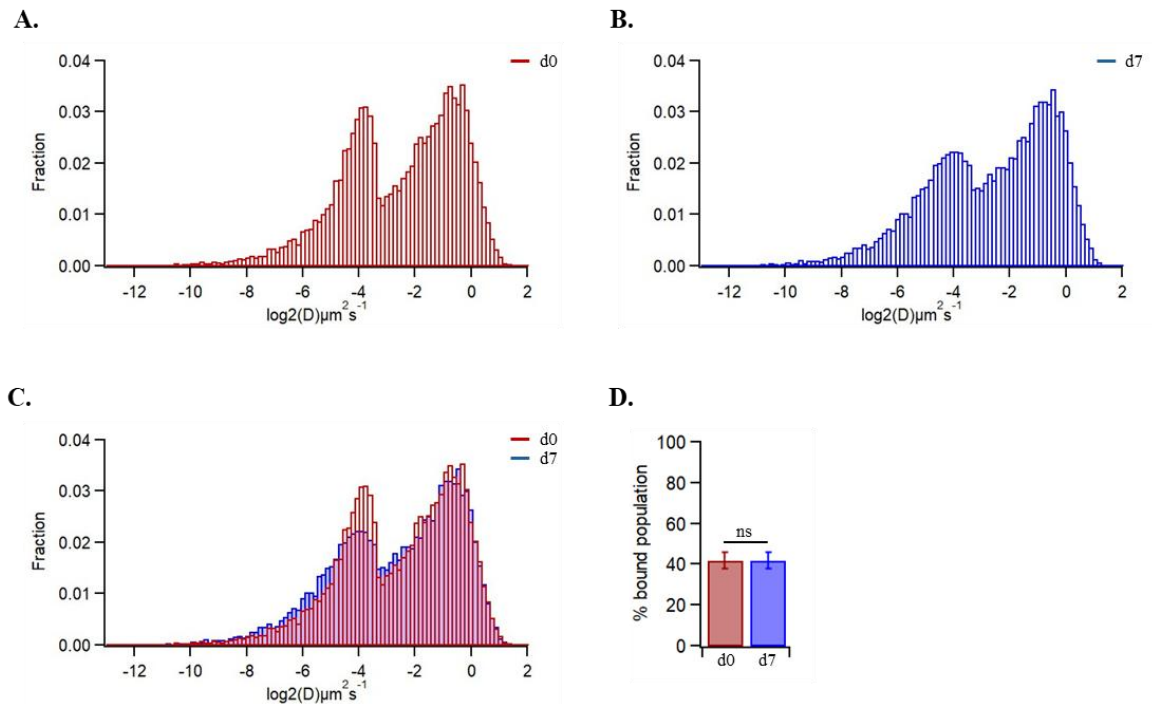


Figure 3.3 – SmcHD1 dynamics across the whole nucleus

Figure 3.3 – SmcHD1 dynamics across the whole nucleus:

A. Histogram of diffusion coefficients of obtained tracks from across the whole nucleus, normalised for the number of analysed tracks ($n = 13591$ from 60 movies). The imaged cells are undifferentiated T7Halo-SmcHD1 R26-eGFP-CIZ1 mESCs, kept for 24h with doxycycline to induce Xist expression. **B.** Histogram of diffusion coefficients of obtained tracks from across the whole nucleus, normalised for the number of analysed tracks ($n = 25735$ from 64 movies). The imaged cells are T7Halo-SmcHD1 R26-eGFP-CIZ1 mESCs which underwent ML differentiation for 7 days. **C.** Comparison of the histograms from A and B corresponding to dynamic behaviour of SmcHD1 at day 0 (red) versus day 7 (blue) mESCs. **D.** Bar plot showing the average percentage of bound molecules of SmcHD1 at day 0 and day 7 differentiated mESCs, from 3 independent replicates. Error bars represent the STDEV.

3.5 Characterization of changes in SmcHD1 dynamic behaviour over the inactive X territory during X chromosome inactivation

To look specifically at SmcHD1 molecules within the Xi domain, I added two key steps to the usual SPT experimental setup (Figure 3.4 A, Section 2.4.1-2). Firstly, taking advantage of the introduced eGFP-CIZ1, a snapshot of the cell is taken before and after every acquisition of a movie containing tracks of single SmcHD1 molecules to account for movement of the Xi domain (Figure 3.4 B). Secondly, at the analysis stage, once the particles have been localized, they are filtered using custom scripts (designed by Dr. Aleksander Szczurek, Dept of Biochemistry, University of Oxford), to extract only the molecules located within the Xi region. In the same way, for each analysed Xi region, molecules located within a randomly selected nuclear region, with the same dimensions as the Xi region are used as a control (Figure 3.4 B, Section 2.4.2).

Using this method, analysis of SmcHD1 dynamics at day 0 within the Xi domain showed an almost equal distribution between the two populations, bound and fast (Figure 3.4 C). In contrast, over the random region, the fast population is larger than the bound population (Figure 3.4 D), similar to the pattern observed over the whole nucleus (Figure 3.3 A). Thus, SmcHD1 has a more prominent bound population over the Xi, about 51%, compared to a random region at day 0, around 37% (Figure 3.4 E, I). At day 7, over the Xi chromosome, the majority of SmcHD1 molecules are bound (Figure 3.4 F), while, in a random region, the fast population is still prevalent, as at day 0 (Figure 3.4 G). The difference in dynamic behaviour between the Xi and the random region is more evident at day 7 compared to day 0, with the protein having a larger bound population in the former, about 60%, than in the latter, around 34%, area (Figure 3.4 H-I). Importantly, within the Xi domain, SmcHD1's bound population increases from day 0 to day 7, as X chromosome silencing progresses (Figure 3.4 I, J). In the random region, the population distribution doesn't change with cellular differentiation, highlighting the specificity of the dynamic changes seen over the Xi as a consequence of chromosome-wide repression (Figure 3.4 K).

To gain a better understanding of how tightly the bound population of SmcHD1 is interacting with the underlying chromatin, I measured the dynamics of histone H2B, which as a constituent of the nucleosome, represents a fully bound state. For this, I randomly integrated a transgene encoding Halo-tagged histone H2B in the interspecific iXist-ChrX XX mESC line (Nesterova et al., 2019) described in section 3.2 (Appendix - Figure A1). By comparing the dynamics of SmcHD1 over the Xi region at day 7 with the dynamics of histone H2B, it can be seen that the bound molecules of SmcHD1 are as immobile as H2B (Figure 3.4 L). The diffusion coefficients values determined for histone H2B are in agreement with those reported in another study that used the same plasmid for cell line generation, but in XY mESCs (Huseyin & Klose, 2021).

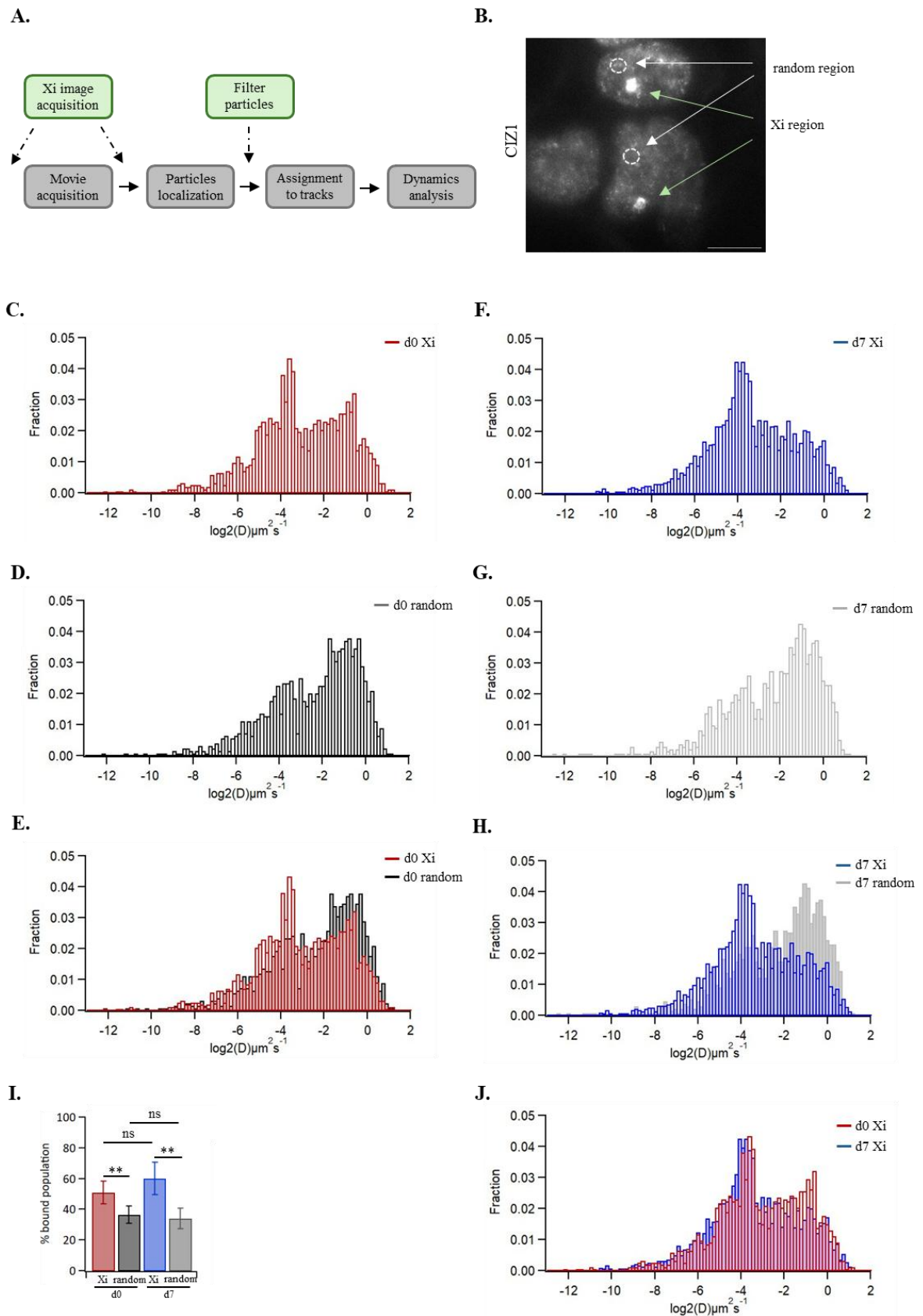


Figure 3.4 – SmcHD1 dynamics across the inactive X territory

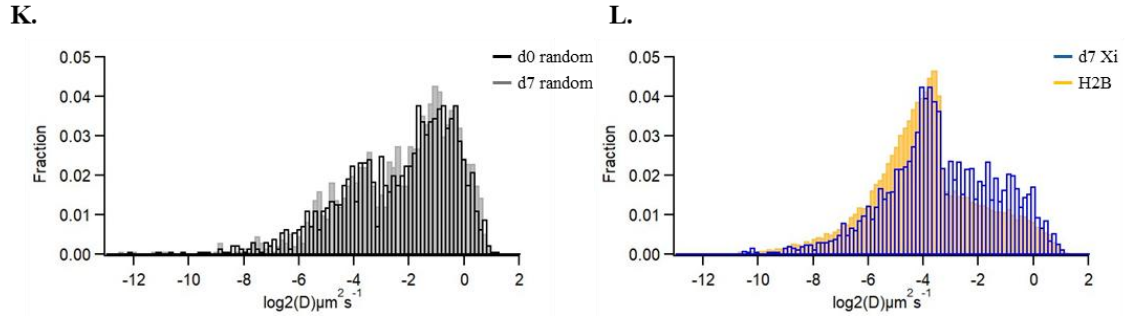


Figure 3.4 (continued) – SmcHD1 dynamics across the inactive X territory:

A. Schematic of Xi-focused SPT analysis. **B.** Representative image of Xi domains marked by CIZ1 accumulation, taken on the TIRF/HILO microscope used for SPT. Two examples of where control random regions can be drawn are also shown. Scale bar corresponds to 10 μ m. **C-D.** Histogram of diffusion coefficients of obtained tracks from across the Xi territory (C) or a random region in the nucleus (D), normalised for the number of analysed tracks ($n = 1521$ from 343 Xi in C and $n = 1238$ from 343 random regions in D). The imaged cells are undifferentiated T7Halo-SmcHD1 R26-eGFP-CIZ1 mESCs, kept for 24h with doxycycline. **E.** Comparison of the histograms C and D corresponding to dynamic behaviour of SmcHD1 at day 0 mESCs over the Xi territory (red) versus a random nuclear region (black). **F-G.** Histogram of diffusion coefficients of obtained tracks from across the Xi territory (F) or a random region in the nucleus (G), normalised for the number of analysed tracks ($n = 2714$ from 307 Xi in F and $n = 1340$ from 307 random regions in G). The imaged cells are T7Halo-SmcHD1 R26-eGFP-CIZ1 mESCs which have underwent ML differentiation for 7 days. **H.** Comparison of the histograms from F and G corresponding to dynamic behaviour of SmcHD1 in day 7 differentiated mESCs over the Xi territory (blue) versus a random nuclear region (grey). **I.** Bar plot showing the average percentage of bound molecules of SmcHD1 at day 0 and day 7 differentiated mESCs over the Xi territory and a random nuclear region, from 5 independent replicates. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test. **J.** Comparison of the histograms from C and F corresponding to dynamic behaviour of SmcHD1 over the Xi territory in day 0 (red) versus day 7 differentiated mESCs (blue). **K.** Comparison of the histograms from D and G corresponding to dynamic behaviour of SmcHD1 over a random region at day 0 (black) versus day 7 differentiated mESCs (grey). **L.** Yellow: histogram of diffusion coefficients of obtained tracks from across the whole nucleus, normalised for the number of analysed tracks ($n = 27547$ from 22 movies). The imaged cells are histone H2BHalo mESCs. Blue: histograms from G corresponding to dynamic behaviour of SmcHD1 over the Xi territory in day 7 differentiated mESCs.

3.6 Discussion

In this chapter, I describe the generation and optimisation of a cell-based model system for SmcHD1 SPT and the development of a methodology that allows analysis of the dynamic behaviour of SmcHD1 molecules in the nucleus and specifically over the Xi chromosome.

Observing SmcHD1 live with the SmcHD1-Snap cell line revealed a preference for accumulation over heterochromatic pericentric regions. While such a localization has been observed previously in a small population of cells using immunofluorescence (Brideau et al., 2015; Ichihara et al., 2022), this is the first report showing this is a population-wide characteristic and it is in agreement with a proteomic study that identified SmcHD1 as an interactor of pericentric chromatin in a pull down in mESCs, using specific probes against major satellites (Saksouk et al., 2014). The difficulty in detecting this localization pattern via immunostaining could be due to either the antibody's epitope being masked when the protein interacts with pericentric regions or SmcHD1 having a highly dynamic interaction with these chromatin locations which is, thus, lost during cell fixation. In support of the latter option, a prior study looking at localization of DNA binding factor MeCP2 using live-cell imaging found that various mutants of this protein localised to pericentric regions, even though immunofluorescence analysis in fixed cells had indicated that they didn't. It was concluded that factors with fast dynamic interactions with chromatin ($<5s$) are inefficiently cross-linked by formaldehyde and are therefore lost during subsequent immunofluorescence sample preparation steps (Schmiedeberg et al., 2009). Moreover, probably for the same reason, all five of the Polycomb group RING finger (PCGF) proteins form distinct foci over the Xi domain in live-cell imaging, but in several cases, enrichment by immunostaining was not observed, despite the same anti-GFP antibody being used in each case (Almeida et al., 2017).

Notably, the observed localization pattern is consistent with prior work showing that the SmcHD1 interacting protein LRIF1 together with HP1 family proteins mediates association with the chromatin mark H3K9me3 which is enriched at pericentric domains (Brideau et al., 2015). This pathway will be explored in detail in the next chapter. Altogether, these considerations highlight the benefit provided by studying SmcHD1 with live-cell imaging.

Since the SmcHD1-Snap cell line was found to be unsuitable for Xi-specific SPT, I generated the T7Halo-SmcHD1 R26-eGFP-CIZ1 line which met all the requirements and was chosen as a model system for this research. The quantification of the live-cell imaging timecourse carried out across 7 days with cells undergoing ML differentiation suggests SmcHD1 gradually accumulates over the Xi chromosome rather than being recruited only later during the initial stages of XCI, as previously inferred from immunofluorescence timecourses (Bowness et al., 2022; Gendrel et al., 2012). The reason behind this difference may be that previous studies have assessed domain formation, so enrichment of SmcHD1 signal over the Xi territory compared to the rest of the nucleus, but this requires a certain number of SmcHD1 molecule to be present over the Xi territory for a domain to become visible by eye. Alternatively, SmcHD1 may only reside for a short time across the Xi in the early stages of XCI, evading detection by fixation-based methods, as discussed above for pericentric regions. In later stages a longer binding time to the underlying Xi chromatin may enable this interaction to be captured also in immunostaining assays. As described in Section 1.3.2.1, the mechanism underlying differentiation-dependent recruitment of SmcHD1 to the Xi chromosome is a long-standing question. Mutations that could perturb this recruitment mechanism and give more insight into its workings will be investigated in Chapter 4.

With the chosen model system, I have initiated an analysis of the dynamic behaviour of SmcHD1 using SPT. Over the whole nucleus, at both analysed cellular states, the observed bound population probably represents SmcHD1 bound at the pericentric regions, autosomal

targets such as the *Pcdh* and *Hox* genes (Benetti et al., 2022; Chen et al., 2015; Gendrel et al., 2013; Mould et al., 2013) and the Xi territory. The newly setup Xi-specific analysis is key for studying changes in SmcHD1 interaction with chromatin over the Xi domain. I show that SmcHD1 has a larger bound population over the Xi compared to a random region and that, within the Xi, this population increases from day 0 to day 7 as XCI progresses. Altogether, this dynamic data is in accordance with SmcHD1 accumulating and binding to the Xi to perform its functions in gene silencing, CpG island methylation and remodelling of X-chromosome architecture (Blewitt et al., 2008; Bowness et al., 2022; Gdula et al., 2019; Gendrel et al., 2012; Gendrel et al., 2013; Jansz, Keniry, et al., 2018; Mould et al., 2013; Nozawa et al., 2013; Wang, Colognori, et al., 2019; Wang et al., 2018). Notably, I can rule out that the dynamics observed within the Xi are an artifact caused by analysing a small region compared to the whole nucleus, since the random region has different dynamics compared to Xi. Nevertheless, such an analysis does reduce drastically the number of analysable tracks compared to the commonly used whole nucleus analysis, which highlights the importance of using a cellular system with robust and synchronous Xist induction to avoid variability coming from varying Xi domain sizes and efficiency of differentiation. Therefore, the inducible Xist setup employed here is key for the Xi-specific analysis.

Notably, in the study mentioned in the introduction of this chapter, looking at RNAPII across the Xi chromosome, the difference in dynamic behaviour of RNAPII between the Xi and nucleoplasm is opposite to that observed for SmcHD1 (Collombet et al., 2023). This is in accordance with the need to shut-down RNAPII-mediated transcription to achieve X-linked gene silencing with the help of many factors, including SmcHD1.

One more point to discuss is that the bound population of SmcHD1 within the Xi at day 0 is perhaps surprisingly large considering the absence of enrichment of the protein at this stage and the need for differentiation for this to take place (Bowness et al., 2022). However,

it should be noted that, after 24h of Xist expression, many of the players involved in XCI are already recruited and, consequently, the chromatin environment has already undergone significant changes, including deposition of H2AK119ub1 which, as previously mentioned, is required for localization of SmcHD1 at the Xi (Jansz, Nesterova, et al., 2018; Loda et al., 2022). The presence of this chromatin mark may play a role in stabilizing SmcHD1 over the Xi territory at day 0, and this will be further explored in the subsequent chapters.

Also, it was interesting to see that overall the diffusion coefficients obtained for SmcHD1 are lower than those of RING1B, which is part of the catalytic core of PRC1, and closer to those of cohesin subunit Scc1 and its loader Scc2 (Huseyin & Klose, 2021; Rhodes et al., 2017). As previously mentioned, PRC1 facilitates gene repression via deposition of the histone modification H2AK119ub1, including during XCI, (de Napoles et al., 2004; Wang et al., 2004), while Scc1 and Scc2 play a key role in chromosomal organisation which, in turn, has been proposed to regulate gene expression through various mechanism such as by affecting the 3D positioning of distal regulatory elements with respect to gene promoters (Horsfield, 2023). It is, thus, tempting to speculate that the proposed role of SmcHD1 in X-linked gene silencing is directly linked to its inferred involvement in altering the architecture of the Xi, rather than for example, by recruitment of other silencing factors. However, more information is required for a better understanding of this question and, thus, the mechanism of action of SmcHD1 will be further dissected using different mutants in Chapter 5.

Chapter 4: Studying the role of LRIF1 in modulating the dynamic behaviour of SmcHD1

4.1 Introduction

The finding from live-cell imaging that SmcHD1 is enriched over the pericentric heterochromatin regions in all cells (Section 3.2) has led me to investigate in more detail the pathway that enables this localization. Pericentric regions are enriched in the repressive histone modification H3K9me3 and, as previously mentioned, LRIF1 (and its human homolog, HBiX1), interacts directly with SmcHD1 to mediate association with H3K9me3-marked chromatin by binding HP1 proteins, that in turn bind H3K9me3 (Brideau et al., 2015; Nozawa et al., 2013). Indeed, enrichment of SmcHD1 at pericentric domains was specifically looked at, but, surprisingly, only observed it in a minority of cells (Brideau et al., 2015; Ichihara et al., 2022). Based on proteomic analysis, it was proposed this may be due to LRIF1 mediating the interaction preferentially with HP1 γ , which is found at H3K9me3-marked locations on chromosome arms, over HP1 α and HP1 β , which bind at pericentric regions (Brideau et al., 2015; Hayakawa et al., 2003; Minc et al., 1999; Minc et al., 2000). However, as discussed in Chapter 3, the use of immunostaining instead of live imaging is probably the reason behind the discrepancy, either due to the antibody's epitope being hidden or due to the dynamic nature of the SmcHD1's interaction with pericentric regions. Moreover, it is worth noting that the population-wide enrichment of SmcHD1 at these heterochromatic domains was observed also for Snap- and not only Halo-tagged

SmcHD1 advocating for this being a bona fide localization pattern rather than an artifact caused by the Halo tag. In this chapter, I will investigate whether LRIF1 is involved in pericentric localization of SmcHD1 as originally hypothesized.

The LRIF1/HBiX1-HP1-H3K9me3 localization pathway has been looked at also in the context of XCI. Localization of SmcHD1 over the Xi territory was not affected by removal of LRIF1/HBiX1 in somatic cells (Brideau et al., 2015; Nozawa et al., 2013), suggesting that, at this developmental stage, SmcHD1 doesn't depend on LRIF1 for Xi enrichment. It is important to mention that the role of H3K9me3 in mouse XCI is still under debate. Unlike in humans and other mammals, the mark is not enriched on the Xi domain using immunofluorescence imaging, contrasting with other repressive histone marks such as H2AK119ub1 and H3K27me3 (de Napoles et al., 2004; Rens et al., 2010). More recent studies, relying on the use of ChIP-seq, have reported the presence of blocks of H3K9me3 over the Xi chromosome, which are enriched compared to the Xa (Ichihara et al., 2022; Keniry et al., 2016) and deposition of this mark by the H3K9 methyltransferase SETDB1 during differentiation of female mESCs was shown to be required for silencing of some of the X-linked genes (Keniry et al., 2016). Importantly, epiblast stem cells (EpiSCs) and MEFs lacking SmcHD1 presented impaired deposition of H3K9me3 across the Xi, regarding the width and height of the blocks. This was accompanied by an expansion of the H3K27me3 blocks that was more evident in MEFs compared to EpiSCs (Ichihara et al., 2022). Although a direct link between erosion of H3K9me3 and SmcHD1-determined gene de-repression could not be established, the authors proposed that SmcHD1-deficient cells fail to correctly establish the H3K9me3 blocks, causing abnormal deposition of H3K27me3 and formation of an Xi which is not condensed enough and, therefore, susceptible to the action of factors that counteract gene silencing (Ichihara et al., 2022). In this context, it would be interesting to explore further the relationship between SmcHD1 and H3K9me3 on the Xi and whether

LRIF1 has any involvement during the establishment phase of XCI. I will address these in my model system in this chapter.

4.2 LRIF1 is required for localization of SmcHD1 to pericentric regions

To investigate the role of LRIF1 in localization of SmcHD1 to pericentric domains, I generated a CRISPR/Cas9-mediated knock-out, using two sgRNAs, of endogenous *Lrif1* in the T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line, described in Chapter 3 (Figure 4.1 A). LRIF1 presents three isoforms, two at a lower molecular weight, of 27kDa and 29kDa, and one at a higher molecular weight, of 85kDa. The deletion of about 13.5kb generated by the sgRNAs-targeted cuts affects the coding sequence of all isoforms and should, thus, prevent their expression. The obtained clones were screened by Western Blot on whole cell lysates and the successful depletion of the lower molecular weight isoforms can be seen in Figure 4.1 B. However, I was not able to detect the 85kDa isoform in WT samples, neither in whole cell lysates nor in nuclear extracts in the soluble fraction which is usually analysed (Figure 5C). The long isoform of LRIF1 could only be found in the insoluble chromatin fraction, suggesting its very tight interaction with chromatin, and, in the LRIF1 KO clone selected for further experiments, it can be seen that also this isoform is efficiently depleted (Figure 4.1 C).

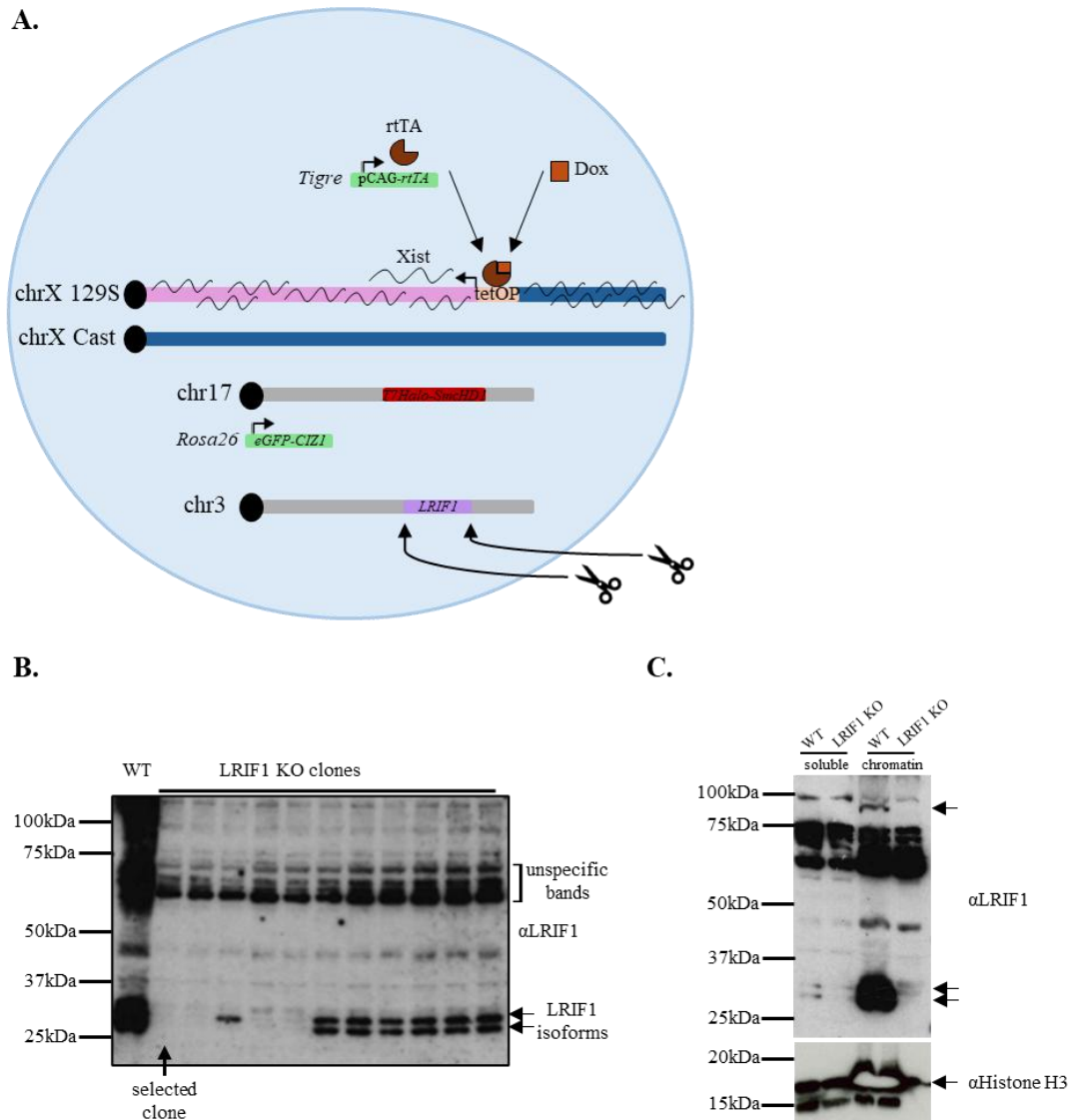


Figure 4.1 – Generation of an *LrIF1*^{-/-} mESC line:

A. Schematic showing the T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line and the generation of an endogenous *LrIF1* knock-out using two sgRNAs. **B.** Western Blot showing the expression of LRIF1 in whole cell lysates from WT mESC (first lane) and from potential *LrIF1*^{-/-} clones obtained (subsequent 11 lanes); the unspecific bands act as a loading control. **C.** Western Blot from nuclear extract showing the soluble fraction in the first two lanes and the insoluble chromatin fraction in the last two lanes. The blot shows the expression of LRIF1 and histone H3. Note that the amount of protein in the chromatin fraction samples could not be quantified due to the nature of the extract, but unspecific bands on the blot suggest there is roughly the same amount loaded between the WT and LRIF1 KO samples.

Having successfully obtained an *Lrif1*^{-/-} mESC line, I went ahead to evaluate the effects of this knockout on SmcHD1 localisation using live-cell imaging. I observed a substantial loss of enrichment of SmcHD1 over pericentric regions across all cells, even though the regions are still clearly present as determined by staining with the vital dye Hoechst (Figure 4.2 A). This confirms that indeed LRIF1 plays a role in SmcHD1 recruitment to pericentric heterochromatin domains. Moreover, the protein level of SmcHD1 was dramatically reduced in the LRIF1 KO condition as shown by Western Blot analysis (Figure 4.2 B), although such a change was not observed at the RNA level (Figure 4.2 C). This is in accordance with the fluorescent images in which, due to the disappearance of the accumulation at pericentric regions, it can be visually appreciated that there is overall less intense SmcHD1 signal (Figure 4.2 A).

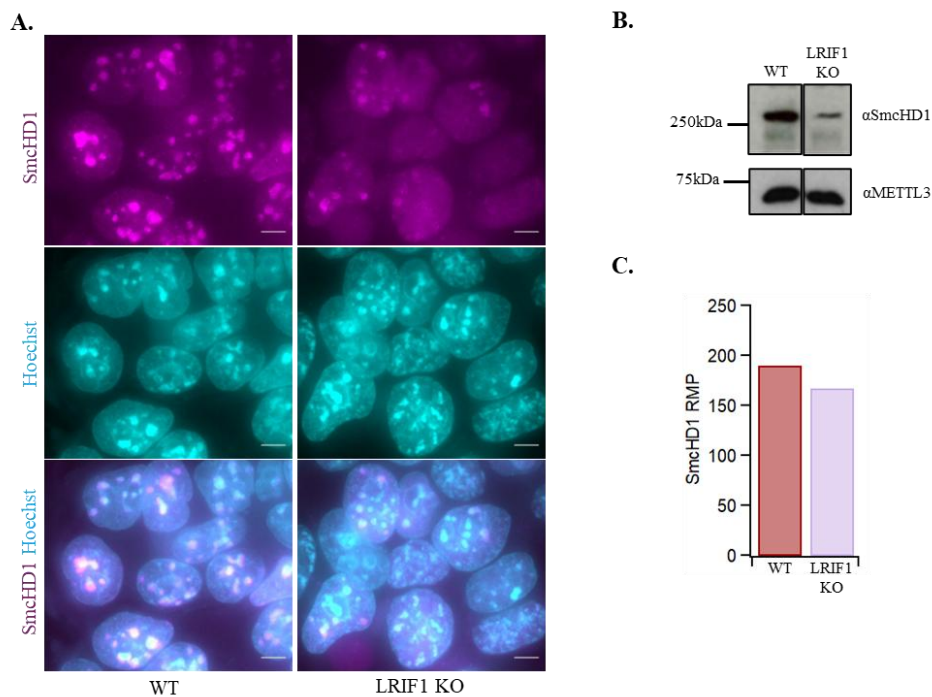


Figure 4.2 – The absence of LRIF1 impairs SmcHD1's localization to pericentric regions:

A. Representative live images of Halo-tag labelling and Hoechst staining in WT and *Lrif1*^{-/-} mESC lines. Images are maximum-intensity projections, scale bar corresponds to 5μm. **B.** Western Blot from nuclear extract showing the expression level of SmcHD1 in the presence and absence of LRIF1; METTL3 acts as a loading control; the extracts were run on the same gel but in non-adjacent lanes. **C.** Relative levels of chromatin-associated SmcHD1 RNA in WT and *Lrif1*^{-/-} mESC lines.

To understand better the reason behind the decrease in protein levels, I analysed SmcHD1 dynamic behaviour using SPT. Regardless of the cellular state, in the absence of LRIF1, SmcHD1 presented a more prominent fast population compared to the bound (Figure 4.3 A-B). Compared to the WT condition, there are indeed more fast-moving molecules (Figure 4.3 A-B), with the amount of bound molecules decreasing from ~42% in WT cells to about 25-28% in cells lacking LRIF1 (Figure 4.3 C-D). This is in accordance with the previously observed loss of SmcHD1, in a subcellular fractionation experiment, from the chromatin-bound fraction in *Lrifi*^{-/-} cells (Brideau et al., 2015) and suggests that unbound SmcHD1 is less stable, leading to the reported overall decrease in its nuclear levels (Figure 4.2 B).

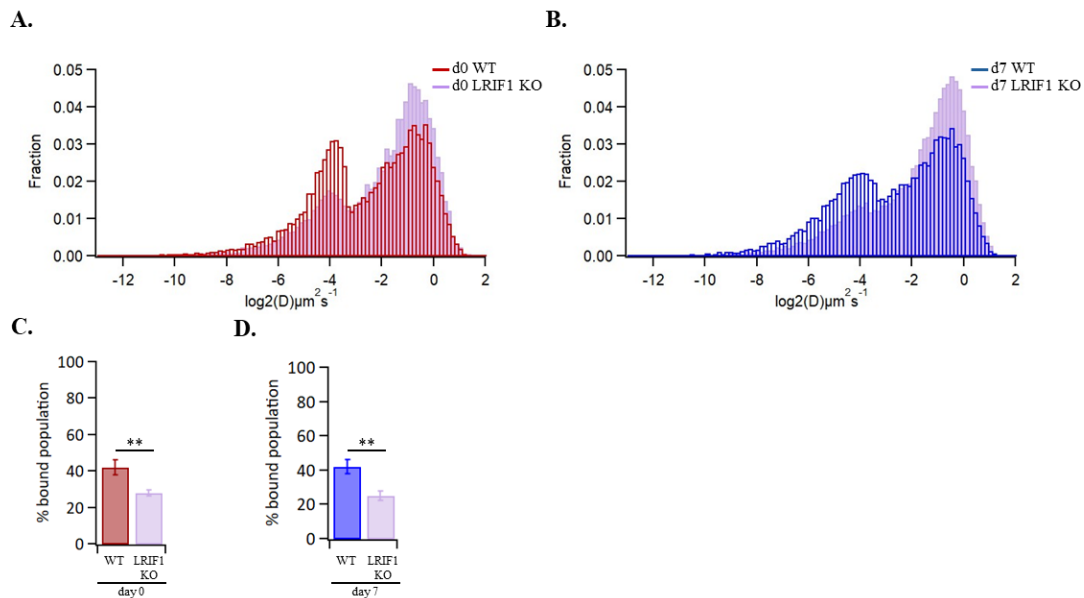


Figure 4.3 – LRIF1’s absence has a striking effect on SmcHD1’s dynamics across the whole nucleus:

A. In purple, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks ($n = 15696$ from 64 movies); the imaged cells are undifferentiated *Lrifi*^{-/-} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **B.** In purple, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks ($n = 27017$ from 66 movies); the imaged cells are *Lrifi*^{-/-} mESCs which have undergone ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **C-D.** Bar plots showing the average percentage of bound molecules of SmcHD1 at day 0 (C) and day 7 differentiated (D) mESCs in the presence and absence of LRIF1, from 3 independent replicates. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

4.3 LRIF1 is required for recruitment of SmcHD1 to the Xi chromosome during ML differentiation

To assess the impact of absence of LRIF1 on SmcHD1 localization to the Xi chromosome, I took the *Lrifi*^{-/-} mESC line through ML differentiation and performed a 7-day live imaging timecourse. Surprisingly, unlike in the WT case, there is no enrichment of SmcHD1 over the Xi domain when LRIF1 is not present (Figure 4.4). Xi accumulation of CIZ1 is not affected indicating that this result is not due to failure to establish Xi. The observed increase in SmcHD1 intensity, particularly at days 3-5, (Figure 4.4 B) is not due to a specific accumulation of the protein over the Xi, but rather from an increase in intensity across the whole nucleus, as can be seen in the images shown in Figure 4.4 A. Using WT cells, previous Western Blot analysis of SmcHD1 levels during NPC differentiation showed no change during the timecourse (Bowness et al., 2022), and, also, based on imaging, I have not observed any such changes in ML differentiation (Figure 4.4 A). More experimental analysis is needed to confirm whether the LRIF1 KO causes SmcHD1 nuclear levels to undergo these fluctuations during differentiation.

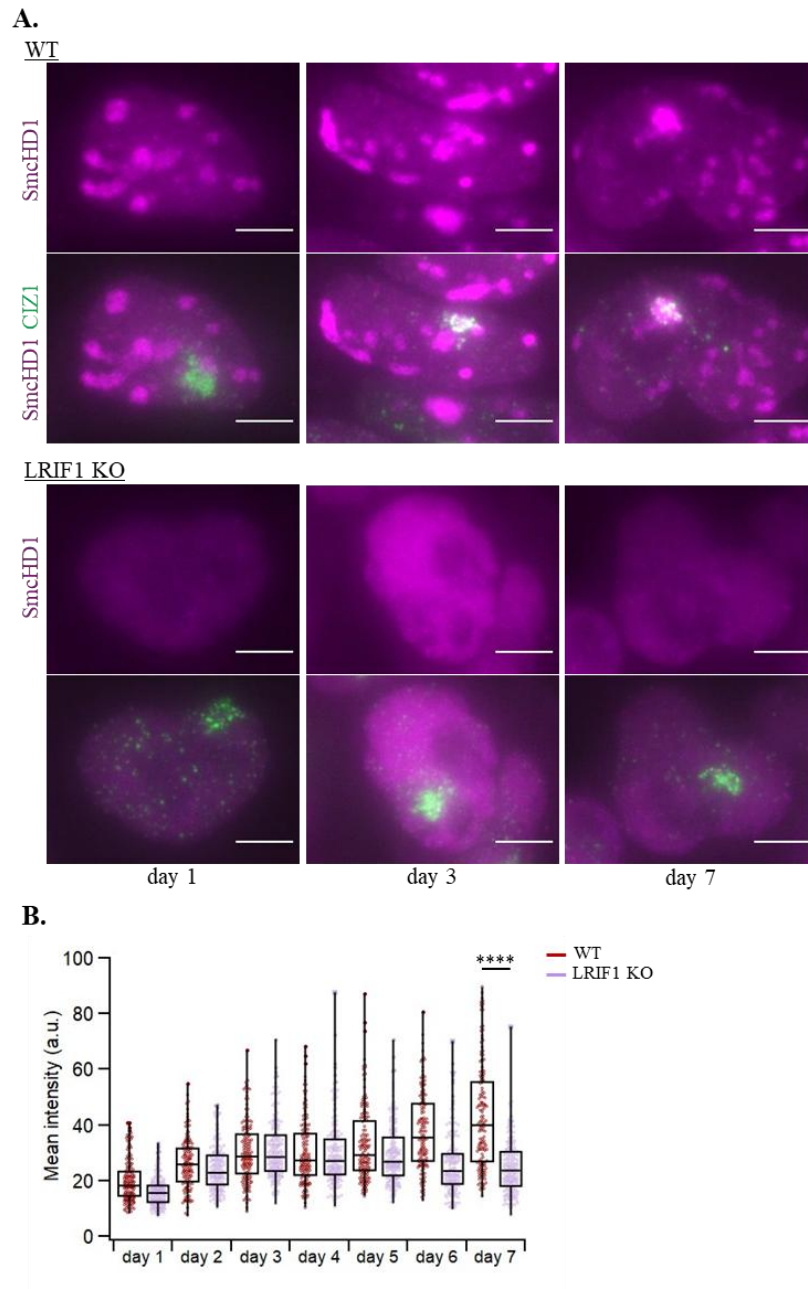


Figure 4.4 - The localization of SmcHD1 to the Xi chromosome is impaired in *Lrifi*^{-/-} cells:

A. Representative live images showing there is gradual enrichment or no enrichment of SmcHD1 over the Xi domain marked by eGFP-CIZ1 from day 1 to day 3 to day 7 of ML differentiation in WT and *Lrifi*^{-/-} mESCs, respectively. Images are maximum-intensity projections, scale bar corresponds to 5µm. **B.** Quantification of SmcHD1 signal underneath the CIZ1 domain, at each day, during a 7 days live imaging ML differentiation timecourse comparing WT (red) and *Lrifi*^{-/-} (purple) mESCs. Each dot represents a cell. n = 136-200. Statistical analysis was done using the unpaired t-test.

I went on to characterize dynamic behaviour of SmcHD1 in the *Lrifi*^{-/-} mESC line, using the Xi-specific SPT methodology described in Chapter 3. Interestingly, already at day 0, removal of LRIF1 greatly diminishes the bound population of SmcHD1 over the Xi territory, from ~51% in WT to ~30% in LRIF1 KO condition (Figure 4.5 A, F). Thus, the specific dynamic behaviour of SmcHD1 over the Xi compared to a random region (Figure 3.4 E) is lost, with both regions displaying a prominent fast population, akin to the profile observed across the whole nucleus (Figure 4.5 B and 4.3 A). More precisely, the bound population within the random region is ~22% (Figure 4.5 F). Similarly, at day 7, in the absence of LRIF1, the bound population of SmcHD1 within the Xi domain decreases from ~60% in WT to ~30% (Figure 4.5 C, F) and the population distributions within the Xi and a random region, where the bound population is ~20%, are extremely similar (Figure 4.5 D, F). Thus, the characteristic change to a larger bound population over the Xi territory which accompanies the silencing of the chromosome (Figure 3.4 J) is completely absent in LRIF1 KO cells, where SmcHD1 molecules retain the predominant fast-moving state from day 0 to day 7 (Figure 4.5 E, F).

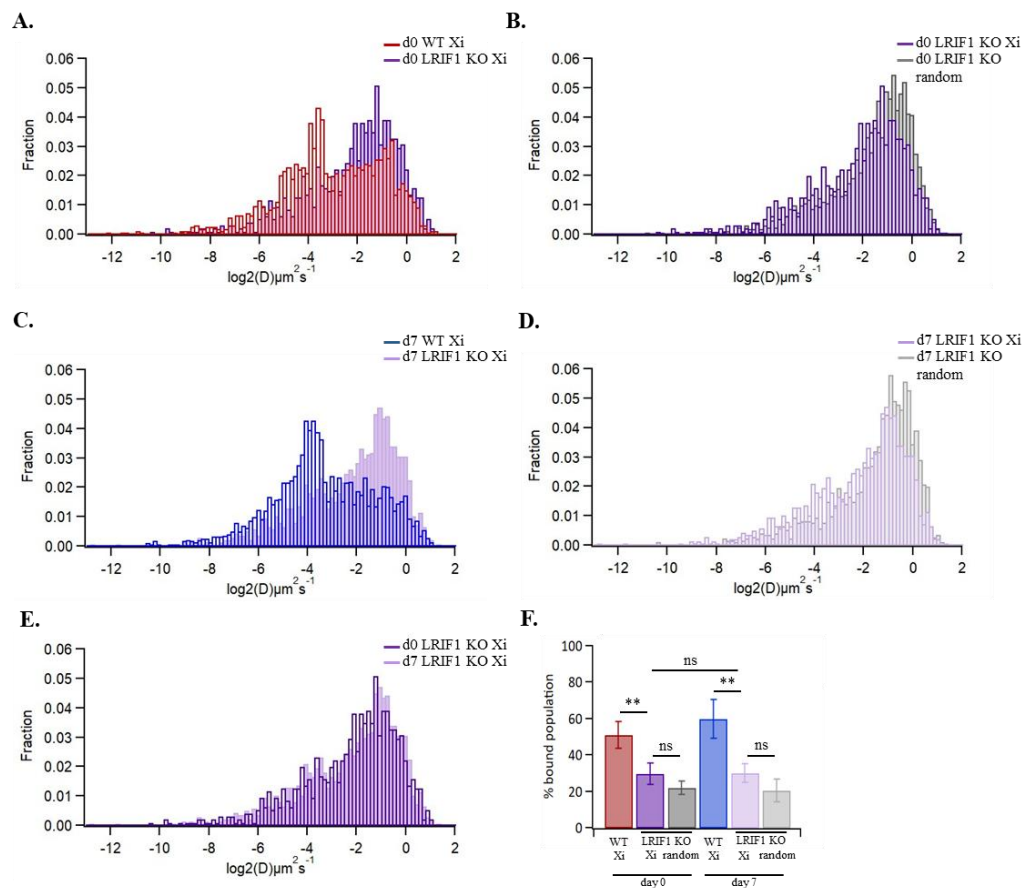


Figure 4.5 - The absence of LRIF1 changes SmcHD1 dynamics across the Xi chromosome

A. In dark purple, histogram of diffusion coefficients of analysed tracks from the Xi territory, normalised for the number of analysed tracks ($n = 942$ from 272 Xi); the imaged cells are undifferentiated *Lrifi*^{-/-} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **B.** In dark grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 1594$ from 272 random regions); the imaged cells are undifferentiated *Lrifi*^{-/-} mESCs, kept for 24h with doxycycline. In dark purple, same histogram as shown in A. **C.** In light purple, histogram of diffusion coefficients of analysed tracks from the Xi territory, normalised for the number of analysed tracks ($n = 1377$ from 235 Xi); the imaged cells are *Lrifi*^{-/-} mESCs which underwent ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **D.** In light grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 1685$ from 235 random regions); the imaged cells are *Lrifi*^{-/-} mESCs which underwent ML differentiation for 7 days. In light purple, same histogram as shown in C. **E.** Comparison of the histograms corresponding to dynamic behaviour of SmcHD1 over the Xi territory at day 0 (dark purple, from A) versus day 7 (light purple, from C) in *Lrifi*^{-/-} mESCs. **F.** Bar plot showing the average percentage of bound molecules of SmcHD1 at day 0 and day 7 differentiated mESCs over the Xi territory and a random nuclear region, in WT (5 independent replicates) and *Lrifi*^{-/-} (4 independent replicates) mESCs. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

To confirm and complement the observation that LRIF1 KO influences the dynamic behaviour of SmcHD1 I performed FRAP assays. FRAP has previously been used to study dynamics of factors involved in X chromosome silencing and revealed how they display altered kinetics over the Xi compared to the nucleoplasm (Almeida et al., 2017; Markaki et al., 2021). FRAP relies on the very fast bleaching of a small region of interest (ROI), using a high intensity laser. The recovery of fluorescence in this region is then followed over time. The images obtained can be used to determine how much of the protein present in the area at the time of bleaching is stably bound and, thus, unable to be replaced by the incoming fluorescent protein molecules. This is similar to the bound population information gained from SPT experiments although it is important to note that the two values may not be the same due to the intrinsic properties of each technique. The key advantage of SPT over FRAP is that it is a concentration-independent assay, while FRAP measurements represent an averaged ensemble behaviour which may not allow the behaviour of some sub-populations to be distinguished and, moreover, do not offer the high temporal resolution that SPT does (Mueller et al., 2013). Importantly, FRAP enables us to calculate how quickly the unbleached fluorescently tagged protein moves into the bleached area i.e. the residence time of the protein of interest in the chromatin environment represented by the bleached region.

A general residence time on chromatin over the whole nucleus can also be calculated from SPT experiments, as previously obtained for other chromatin binding factors, by changing the acquisition settings: photoactivate molecules only before movie acquisition and then follow how long they remain bound for, using low excitation laser intensity and slower frame rate (Huseyin & Klose, 2021; Rhodes et al., 2017). However, despite many trials, I was not able to estimate the residence time of SmcHD1 using this method, due to the residency time being very high and, thus, at the measurable limit imposed by photobleaching. For this reason, the use of FRAP is key to obtain this information.

Furthermore, FRAP will allow me to determine the residence time both over the Xi and a random nucleoplasmic region. This is a separation I would not have been able to obtain with the Xi-specific SPT methodology due to low number of analysable molecules stemming from the acquisition mode itself, coupled with the filtering required for the region-focused analysis.

I conducted FRAP assays in the same cell lines expressing Halo-tagged SmcHD1 used for SPT experiments to avoid confounding results due to different tags potentially altering the behaviour of the protein of interest in different ways. The disadvantage of using Halo-tag over an intrinsically fluorescent tag, like GFP, is that any protein synthesized after labelling with Halo ligand is not fluorescent and, thus, the values obtained for the recovery time and the immobile population may be slightly overestimated. Nevertheless, this will affect the WT and all analysed mutant cell lines to the same extent so comparison across them is viable. Initial trials indicated a very slow recovery, particularly over the Xi domain, which is the main ROI, of WT SmcHD1 as well as other SmcHD1 mutants (which will be shown in Chapter 5). To accommodate this, I increased the experimental timeframe to 20 minutes post bleaching, adjusted the imaging setup to acquire z-stacks to mitigate the movement of the cells over such a long period of time and kept the laser power as minimal as possible to avoid photobleaching and phototoxicity. The setup was kept across all cell lines analysed (both in this chapter and in Chapter 5) to be able to directly compare all of them. The apparent bleaching level achieved across the different analysed conditions varies due to changes in the mobility of the protein which will, thus, populate more or less the bleached area in the lag time (~3.3 seconds) between the bleach pulse and full acquisition of the first z-stack.

Looking at WT SmcHD1, at day 7, snapshots from the FRAP assay show that fluorescence recovers to almost pre-bleach levels in the nucleoplasm, but not on the Xi

domain (Figure 4.6 A). The average fluorescence recovery curves show the turnover for both of these regions is on a timescale of minutes (Figure 4.6 C-D). In the nucleoplasm, it takes about 3 minutes to recover almost 80% of the pre-bleach fluorescence level and, until the end of the experimental timeframe, the recovery is almost complete showing that the immobile population in this genomic context is quite low (Figure 4.6 D). In the Xi territory, the recovery rate is much slower, reaching, for example, only about 15% of the initial fluorescence level in 3 minutes, and the immobile population is much larger (Figure 4.6 C). In contrast, in the absence of LRIF1, the difference in dynamic behaviour between the Xi chromosome and the nucleoplasm is lost, with FRAP snapshots and average fluorescence recovery curves showing fluorescence recovery with similar rates and reaching almost pre-bleach levels in both areas (Figure 4.6 B-D). The turnover of SmcHD1 is, in this case, on a timescale of seconds, with the fluorescence reaching around 80% of the original intensity in about 1 minute (Figure 4.6 C-D). Since in the *Lrifi*^{-/-} cells SmcHD1 does not enrich over the Xi domain as in WT cells, the identification and selection of the ROI to be bleached was achieved using the eGFP-CIZ signal; also, a snapshot in this channel was taken post-bleach, after acquisition of the first z-stack, to confirm the right area has been bleached (this was done also for other mutant cell lines analysed in Chapter 5, where SmcHD1 accumulation is impaired). A further point is that, due to the higher resolution and sensitivity of the spinning-disk confocal microscope used for the FRAP assay, I was able to discern some SmcHD1 accumulation over the Xi chromosome in a minority of the cells, an example of which can be seen in Figure 4.6 B. However, the enrichment was not as strong as that observed in WT cells.

For a more quantitative comparison between the two cell lines, I fitted the FRAP curves with a single exponential function. Usually, such curves for chromatin binding factors are fitted with a biexponential function which reflects the two average modes of binding that

such proteins normally possess: firstly, more transient binding events which will manifest with a fast recovery rate upon photobleaching and secondly, tighter binding events which will display a slow recovery rate. As seen in the SPT data, SmcHD1 also appears to exhibit this behaviour. However, across the Xi domain, at day 7, the bound population is very prominent and, accordingly, upon fitting the corresponding FRAP curve with a biexponential function, the lifetimes of the two states were equal (~ 6 min), suggesting that the contribution of the transiently binding population is negligible. In the *Lrifi*^{-/-} cells, the recovery was so quick both within the Xi territory and the nucleoplasm that inferring information about the fast-recovering molecules depends solely on the first datapoint, making a biexponential fit unreliable. For similar reasons, this fit is not suitable for other analysed mutants either, which will be shown and discussed in Chapter 5. Taking all this in consideration, I opted for the single exponential fitting (Figure 4.6 E-G).

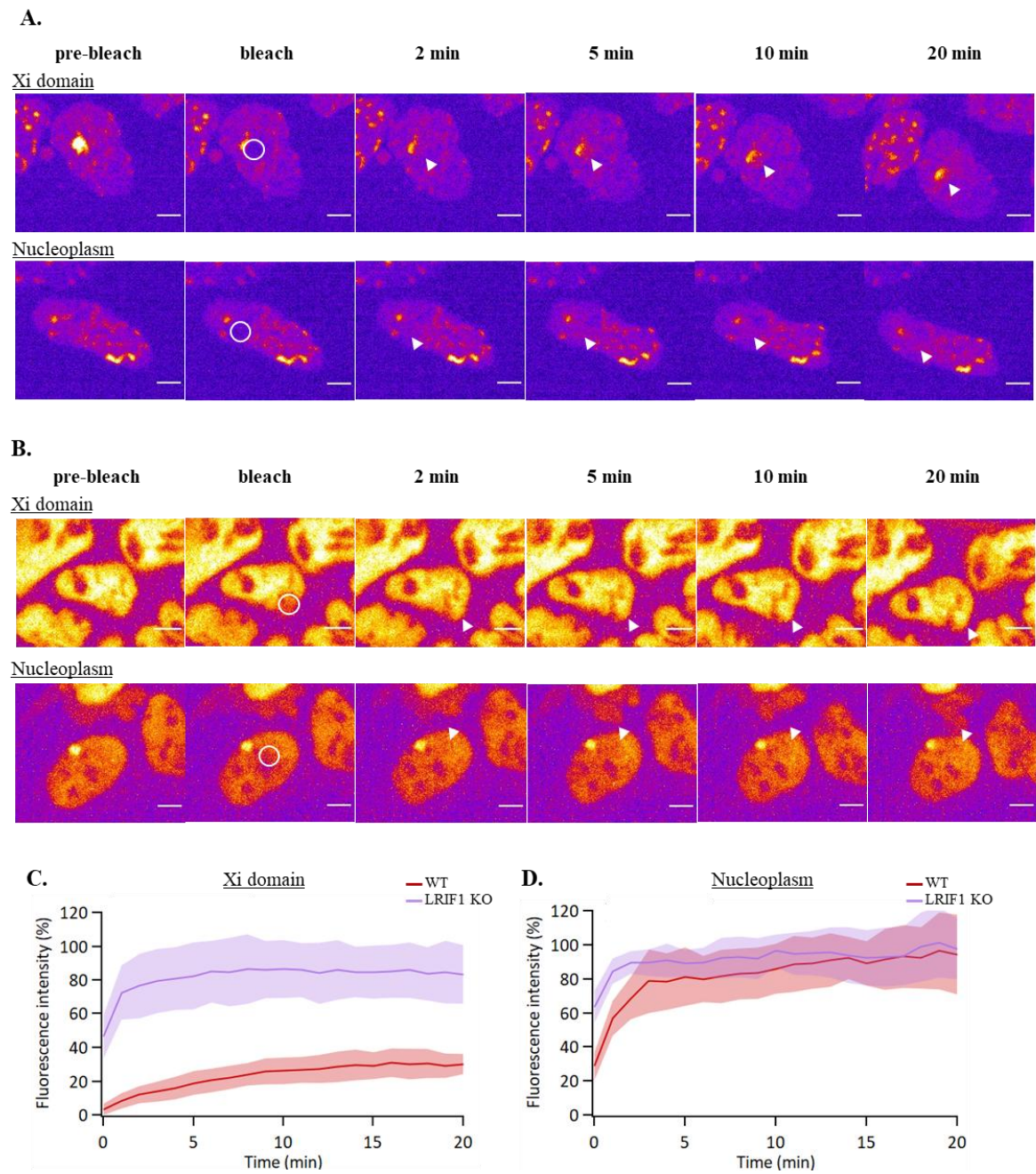


Figure 4.6 – The effect of LRIF1's absence on SmcHD1's dynamics assessed by FRAP

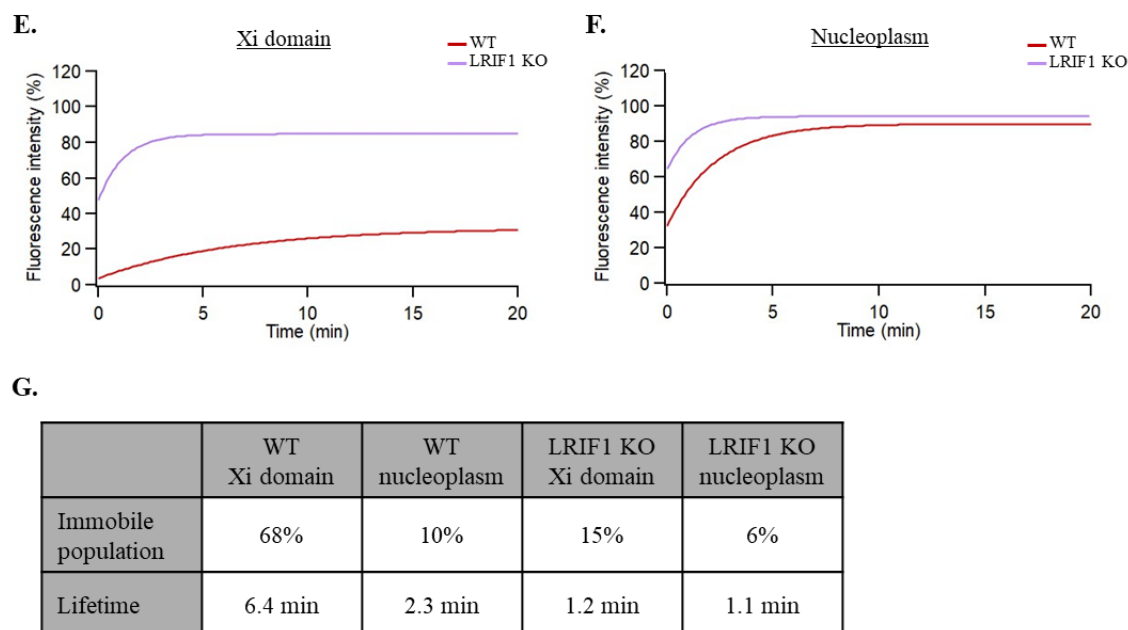
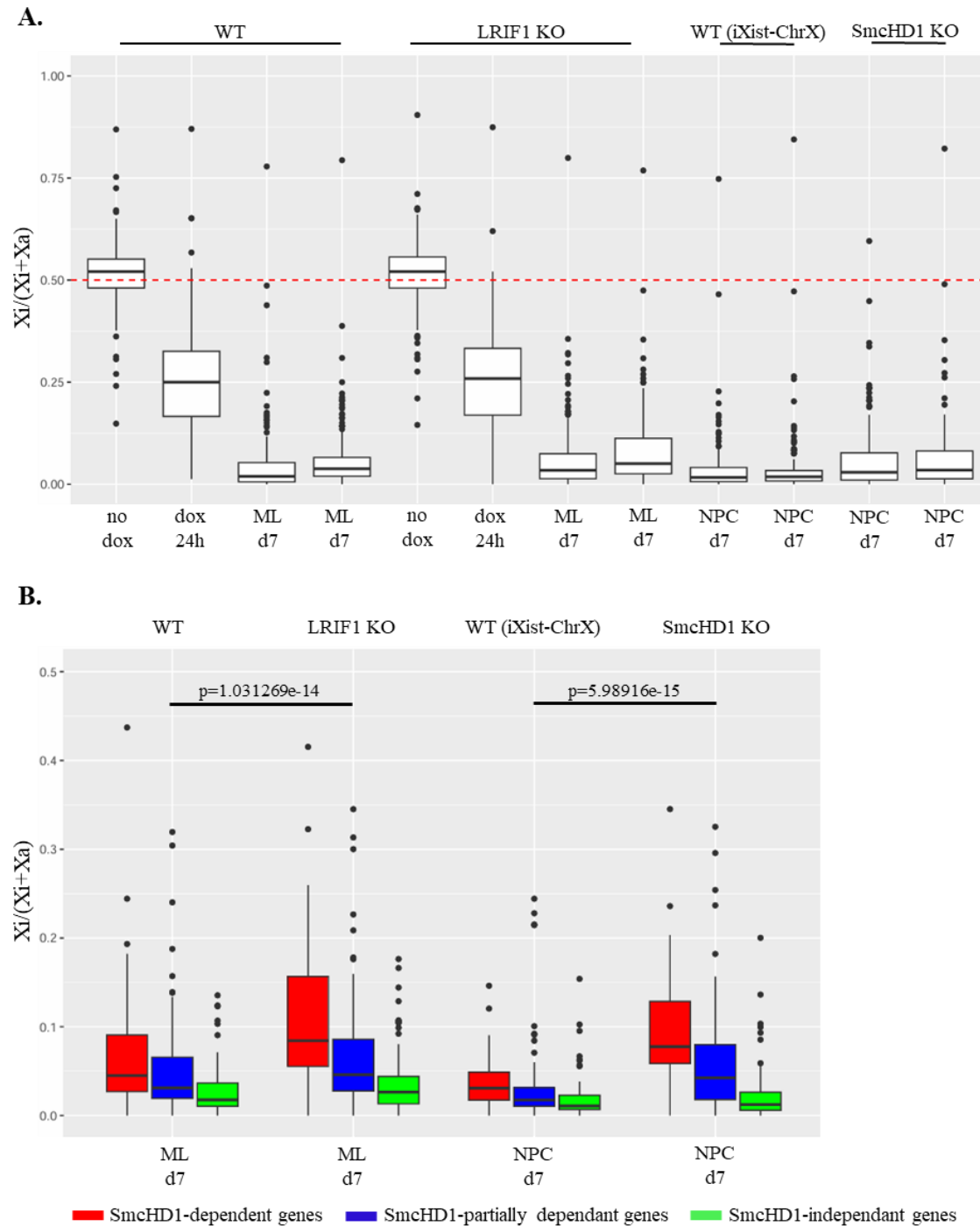


Figure 4.6 (continued) – The effect of LRIF1 absence on SmcHD1 dynamics assessed by FRAP:

A-B. Representative images from a FRAP experiment using the WT (A) or *Lrifi*^{-/-} (B) mESC which underwent ML differentiation for 7 days. The bleached region is marked by a white circle on the bleach snapshot and by a white arrow in the post-bleach snapshots. Scale bar corresponds to 5µm. **C-D.** Average fluorescence recovery curves over the Xi domain (C) or a random nucleoplasmic region (D) comparing WT and *Lrifi*^{-/-} mESCs which underwent ML differentiation for 7 days. The shading represents the STDEV. 16, 15, 21 and 15 cells were analysed for Xi domain WT, Xi domain LRIF1 KO, nucleoplasm WT and nucleoplasm LRIF1 KO, respectively. **E-F.** Single exponential fits of the average fluorescence recovery curves over the Xi domain (E) or a random nucleoplasmic region (F). **G.** Summary of the immobile population and lifetime values extracted from the fits shown in E-F.

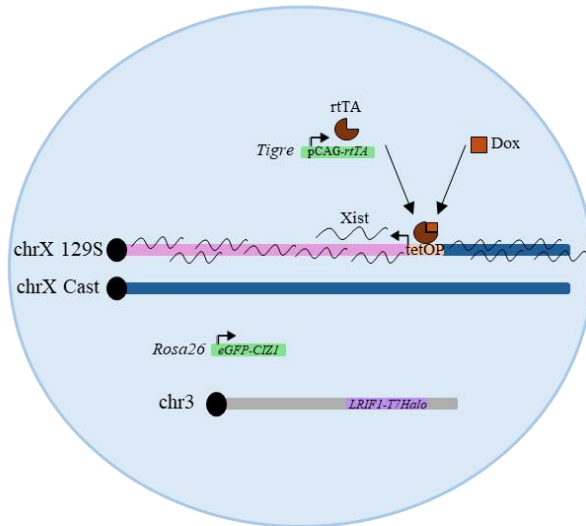
It can then be appreciated that, while the lifetime of SmcHD1 over the Xi domain in WT conditions is almost three times larger than in the nucleoplasm, this difference is completely erased by the absence of LRIF1 and the lifetime is even shorter. Although greatly diminished, a very small difference remains in the immobile populations between the two areas, in the *Lrifi*^{-/-} cells, perhaps due to the Xi being a more condensed area which slows down the movement of the SmcHD1 molecules, even if they are not specifically binding anymore in this condition (Figure 4.6 G).

Seeing the altered SmcHD1 dynamics in the *Lrifi*^{-/-} mESC line, I wanted to investigate whether this has an effect on X-linked gene silencing. To quantify this, I used chrRNA-seq and the calculation of the allelic ratio for X-linked genes, specifically the ratio between the number of reads coming from the 129S allele and the sum of the reads coming from both alleles, which is possible due to the presence of SNPs as described in section 3.2. Looking at all informative genes across the X chromosome, both WT and LRIF1 KO cells showed, in uninduced conditions, the expected allelic ratio ~0.5, which, following 24h doxycycline treatment, dropped to ~0.25 (Figure 4.7 A). At day 7, when SmcHD1 is stably enriched over the Xi domain, gene repression is almost complete in WT cells. However, in the absence of LRIF1, the allelic ratio slightly skewed away from 0 (Figure 4.7 A). To characterize this in more detail, further analysis split genes into three categories according to their dependency on SmcHD1 to achieve full silencing, as previously described (Gdula et al., 2019). This revealed that the genes whose silencing is impaired in *Lrifi*^{-/-} cells are SmcHD1-dependent and, to a lesser extent, SmcHD1-partially dependent genes (Figure 4.7 B). The observed defect is similar to that previously seen in SmcHD1 KO cells taken through NPC differentiation (Figure 4.7 B) (Bowness et al., 2022).

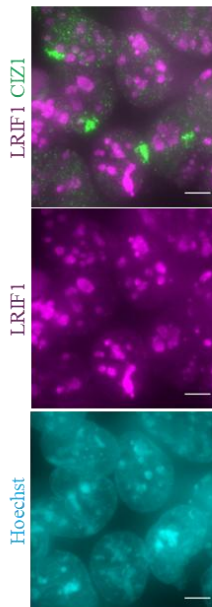


Lastly, given the differences in LRIF1's involvement in SmcHD1 enrichment over the Xi territory between XCI establishment, shown above, and maintenance (Brideau et al., 2015; Nozawa et al., 2013), I decided to look briefly also at the localization relationship between LRIF1 itself and the Xi. For this, I took advantage of an mESC line available in the lab, which, in the parental cell line iXist-ChrX XX (section 3.2), had introduced eGFP-CIZ1 in the *Rosa26* locus, as is the case for my cell lines, and a T7-Halo-tag at the C-terminus of endogenous *Lrifi1*, so that all three isoforms are labelled (Figure 4.8 A). Halo-tag labelling of these cells at day 0 showed LRIF1 has a similar distribution to SmcHD1, being enriched at pericentric regions, but, not over the Xi domain, marked by eGFP-CIZ1 (Figure 4.8 B). After 7 days of ML differentiation, LRIF1 becomes enriched also across the Xi chromosome (Figure 4.8 C), as previously seen in mouse and human somatic cells (Brideau et al., 2015; Nozawa et al., 2013).

A.



B.



C.

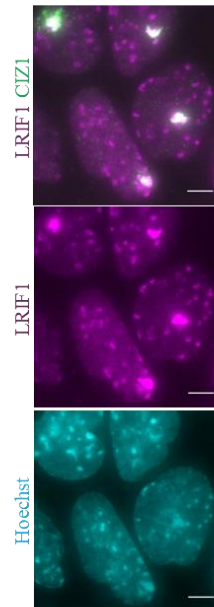


Figure 4.8 – Localization of LRIF1 over the Xi chromosome:

A. Schematic showing the LRIF1-T7-Halo mESC line which has a T7-Halo-tag introduced at the C-terminus of endogenous *Lrifi* and a transgene encoding eGFP-tagged CIZ1 into the *Rosa26* locus. **B-C.** Representative live-cell images of Halo-tag labelling, eGFP-CIZ1 signal and Hoechst staining in LRIF1-T7-Halo mESC line, taken after 24h doxycycline addition (B) or after 7 days of ML differentiation (C). Images are maximum-intensity projections, scale bar corresponds to 5 μ m.

4.4 The interplay between LRIF1-mediated and H2AK119ub1-dependent SmcHD1 recruitment pathways

As described in Chapter 1, recruitment of SmcHD1 to the Xi chromosome has been shown to depend on deposition of the H2AK119ub1 mark by PRC1, via hnRNPK which binds the Xist B/C-repeat (Almeida et al., 2017; Jansz, Nesterova, et al., 2018; Pintacuda et al., 2017). Given the novel involvement of LRIF1 in the recruitment of SmcHD1 presented in this chapter, I wanted to understand better the relationship between these two localization pathways in my model system: are both required? if yes, how are they connected?

As an initial check as to whether H2AK119ub1 influences localization of SmcHD1 to Xi during establishment of XCI in ML differentiation, I relied on the treatment with the proteasome inhibitor MG132 which causes global depletion of H2AK119ub1 (Tavares et al., 2012). I treated the cells for 6 hours with 10 μ M MG132 on the 7th day of differentiation and imaged them immediately after the end of the treatment, as shown in the schematic in Figure 4.9 A. The successful depletion of H2AK119ub1, in two independent replicates, was confirmed by Western Blot (Figure 4.9 B). Imaging of MG132-treated cells showed impaired enrichment of SmcHD1 over the Xi domain, with the percentage of cells having such a domain decreasing from almost 100% to ~40% (Figure 4.9 C-D). However, the drug was not equally efficient across all cells, with some partially affected or unaffected domains also visible, and there were secondary effects due to MG132 treatment, as can be seen from the changed morphology of the nuclei (Figure 4.9 C). Nevertheless, since this result suggests H2AK119ub1 is required for maintenance of SmcHD1 enrichment over the Xi chromosome in ML differentiation, I went on to generate a cell line that would allow for a more in-depth investigation.

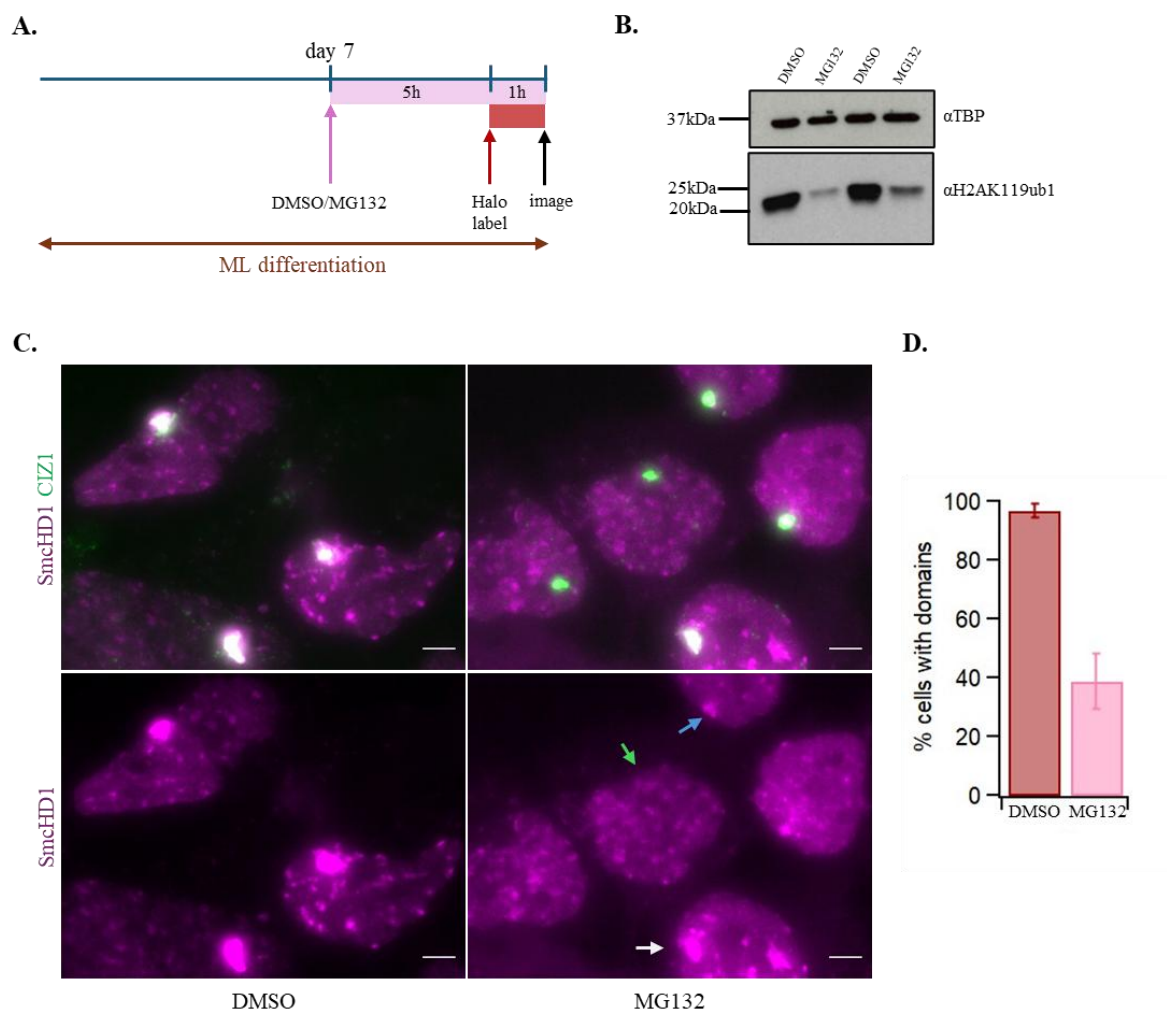


Figure 4.9 – Effect of MG132-induced H2AK119ub1 depletion on SmcHD1 localization to Xi:

A. Schematic showing timing of MG132 treatment in relation to cellular differentiation and Halo-tag labelling and imaging. **B.** Western Blot from RIPA extracts showing the level of H2AK119ub1 upon treatment with MG132 10μM; TBP acts as a loading control. **C.** Representative live images of Halo-tag labelling and eGFP-CIZ1 signal in WT T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line, taken after 7 days of ML differentiation and treatment with DMSO/MG132 10μM. In the bottom right panel, the arrows show varying effects of the MG132 treatment on SmcHD1 enrichment over the Xi domain: completely affected (green), partially affected (blue) and not affected (white). Images are maximum-intensity projections, scale bar corresponds to 5μm. Note: the contrast is not set the same between the DMSO and MG132 conditions due to greatly reduced signal in MG132 condition which didn't allow for a setting that could work for both. **D.** Bar plot showing percentage of cells with SmcHD1 enrichment over the Xi domain at day 7 upon DMSO/MG132 treatment. The average comes from two independent replicates, with around 50 cells analysed in each of them. Error bars represent the STDEV.

Using the WT T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line as a background, I used CRISPR/Cas9-mediated homologous recombination to delete the PID region from the *Xist* locus. The PID region includes the B-repeat and almost all of C-repeat and is required for recruitment of PRC1 to the Xi chromosome (Figure 4.10 A) (Almeida et al., 2017; Pintacuda et al., 2017). Successful removal of the PID region was initially screened for by PCR, looking for appearance of a shorter band of 1.7kb, instead of 3.5kb present in WT (Figure 4.10 B). Then, immunostaining for H2AK119ub1 was carried out in cells which had been treated with doxycycline for 24h, looking for clones which retain CIZ1 domain formation, as its recruitment to the Xi depends on the E-repeat of *Xist* (Ridings-Figueroa et al., 2017; Sunwoo et al., 2017), but lack H2AK119ub1 domains (Figure 4.10 C-D). The clone shown in Figure 4.10 C-D was selected for further experiments.

At day 0, as expected, removal of the PID region did not affect in any way the localization pattern of SmcHD1 over the whole nucleus or specifically over the future Xi territory (Figure 4.10 E, G). After 7 days of ML differentiation, in *Xist*^{ΔPID} cells, enrichment of SmcHD1 at pericentric regions is retained, but it is lost over the Xi domain, as confirmed both by visual inspection of the cells and quantification of the intensity of the SmcHD1 signal under the CIZ1 domain (Figure 4.10 F-G). Due to time constraints I was not able to investigate SmcHD1 dynamic behaviour in the *Xist*^{ΔPID} cell line.

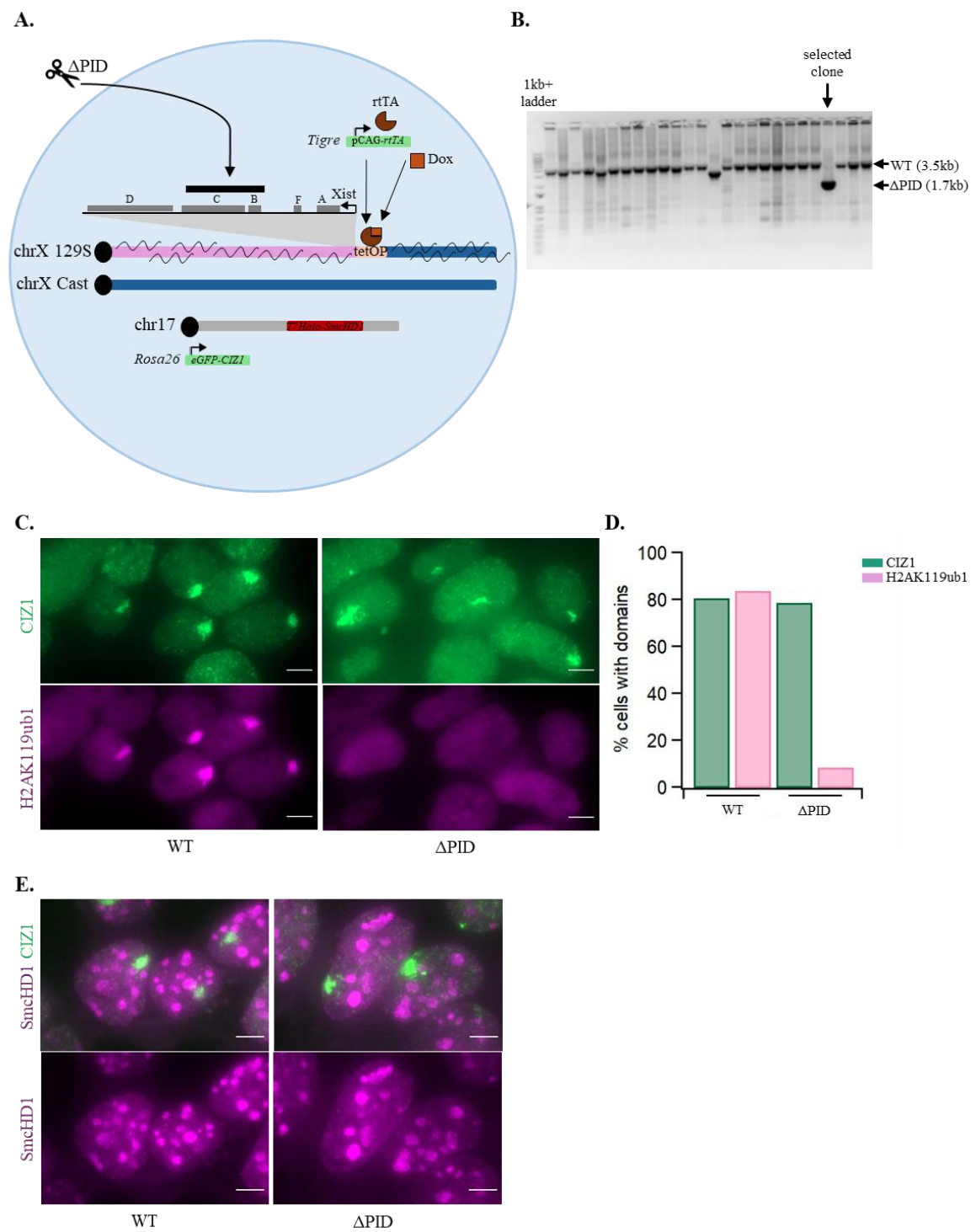


Figure 4.10 – The deletion of Xist PID region affects localization of SmcHD1 to Xi

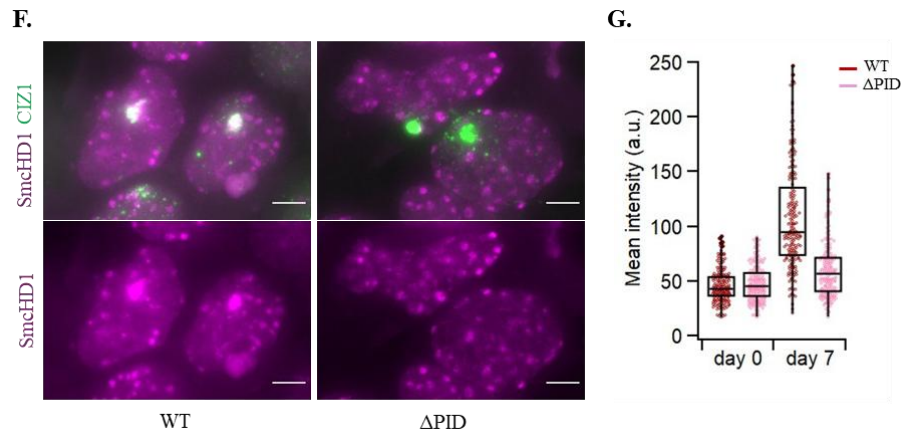


Figure 4.10 (continued) – The deletion of Xist PID region affects localization of SmcHD1 to Xi:

A. Schematic showing the T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line and the generation of a deletion of the PID region of Xist at the endogenous locus. **B.** PCR on gDNA subsequently run on an EtBr agarose gel to screen for deletion of PID region of Xist. **C.** Immunostaining for H2AK119ub1 and eGFP-CIZ1 signal in WT and Xist^{ΔPID} mESC lines, kept for 24h with doxycycline to induce Xist expression. Images are maximum-intensity projections, scale bar corresponds to 5μm. **D.** Bar plot showing percentage of cells with CIZ1 and H2AK119ub1 domains in WT and Xist^{ΔPID} mESC lines, following 24h doxycycline induction. n= 150 cells. **E-F.** Representative live images of Halo-tag labelling and eGFP-CIZ1 signal in WT and Xist^{ΔPID} mESC lines, following 24h doxycycline addition (D) or ML differentiation for 7 days (F). Images are maximum-intensity projections, scale bar corresponds to 5μm. **G.** Quantification of SmcHD1 signal underlying the CIZ1 domain, at day 0 and day 7, comparing WT (red) and Xist^{ΔPID} (pink) mESCs. Each dot represents a cell. n = 182-200.

4.5 Discussion

In this chapter, I evaluated the role of LRIF1 in localization of SmcHD1 to pericentric regions and the Xi chromosome during XCI establishment. This was done through visual inspection of phenotypic changes, assessment of SmcHD1's altered dynamic behaviour, by SPT, FRAP, and quantification of the degree of X-linked gene silencing, all in a context in which LRIF1 was absent. Furthermore, I generated a cell line that will enable a clearer

understanding of the H2AK119ub1-dependent recruitment pathway of SmcHD1 to the Xi and how it relates to the newly observed involvement of LRIF1 in this localization.

For these experiments I generated an *Lrifi*^{-/-} mESC line by deleting a large portion of the endogenous *Lrifi* locus and, thus, perturbing expression of its three isoforms (Figure 4.1). This was done in the background of my main WT cell line, which has a Halo-tagged endogenous SmcHD1 and eGFP-tagged transgenic CIZ1. Imaging of the cells with LRIF1 KO showed severely impacted enrichment of SmcHD1 at pericentric regions (Figure 4.2 A), confirming the previously hypothesized role of the LRIF1-HP1-H3K9me3 pathway in such a localization (Brideau et al., 2015). This is further supported by the observed population-wide accumulation of LRIF1 at these heterochromatic domains (Figure 4.8 B). Such a localization pattern for LRIF1 was reported previously, although in a lower proportion of cells, probably due to the use of antibody-based staining instead of live imaging (Ichihara et al., 2022). The absence of LRIF1 has a striking effect on the dynamic behaviour of SmcHD1, with most of the protein now found in a fast-moving state (Figure 4.3). The profile is similar to that observed when assessing the kinetics of Halo-tag alone, which does not have any specific interaction with chromatin (Huseyin & Klose, 2021). This underlies the extent to which LRIF1 KO removes SmcHD1 from chromatin and suggests that the nuclear-wide bound population of SmcHD1 is almost entirely represented by SmcHD1 interacting with H3K9me3-marked heterochromatin regions.

Intriguingly, in a recent study by Ichihara et al. (2022), it was reported that LRIF1 loses its pericentric localization in EpiSCs and MEFs with an SmcHD1 null mutation, and this is accompanied by a decrease in LRIF1 protein levels, as observed herein for SmcHD1 protein levels (Figure 4.2 B). Together with the data shown in this study, this observation suggests that SmcHD1 and LRIF1 require each other for a stable interaction with pericentric domains: LRIF1 is needed to mediate the interaction between SmcHD1 and HP1-H3K9me3, but

association of SmcHD1 with LRIF1 stabilizes LRIF1's own interaction with these heterochromatic regions. To further explore this potential interdependency, one could remove SmcHD1 in the LRIF1-T7-Halo mESC line and observe whether the nuclear distribution pattern of LRIF1 is affected. It is also worth pointing out that a minority of *Lrifi*^{-/-} cells, retain enrichment of SmcHD1 at some pericentric regions (Figure 4.2 A), which suggests there may be other factors that can facilitate SmcHD1's interaction with these domains, potentially in a cell cycle dependent manner. Identifying such factors would help shed more light on the relationship between SmcHD1 and H3K9me3-marked chromatin, which, as I will discuss below, is of interest also in the context of XCI.

During ML differentiation of *Lrifi*^{-/-} mESCs, there is no gradual accumulation of SmcHD1 over the Xi domain as observed in WT cells (Figure 4.4). This outcome is associated with a prominent loss of the bound population of SmcHD1 at day 7 (Figure 4.5 C, F), leading to loss of the Xi-specific dynamic behaviour compared to a random nucleoplasmic region, as assessed both by SPT (Figure 4.5 D, F) and FRAP (Figure 4.6 B-D). The average fluorescence recovery curves, in the absence of LRIF1, are reminiscent of proteins such as PCGF1, which is not involved in XCI, contrasting with PCGF3/5 which is crucial for Polycomb recruitment to the Xi and displays FRAP curves more similar to WT SmcHD1 (Almeida et al., 2017). All together, these observations highlight a role of LRIF1 in recruitment of SmcHD1 to the Xi chromosome, during the establishment phase of XCI, which is further supported by the SmcHD1 KO-like defect in X-linked gene silencing displayed by *Lrifi*^{-/-} cells (Figure 4.7). Since LRIF1 mediates association of SmcHD1 with H3K9me3, this newly observed role of LRIF1 is interesting also in the context of the recently proposed role of SmcHD1 of ensuring proper deposition of the H3K9me3 blocks across the Xi chromosome (Ichihara et al., 2022).

Notably, as previously mentioned, earlier studies reported removal of LRIF1/HBiX1 has no effect on SmcHD1 localization over the Xi domain in somatic cells where XCI has already been fully established (Brideau et al., 2015; Nozawa et al., 2013). It may be that, while LRIF1 is needed early on to mediate the interaction between SmcHD1 and H3K9me3 and, thus, anchor SmcHD1 to the Xi, as silencing progresses, another pathway/s takes over. In this respect, it would be interesting to explore in more detail the relationship between this localization pathway and the H2AK119ub1-dependent pathway, which was confirmed in this study to be required for accumulation of SmcHD1 over the Xi in mESCs undergoing ML differentiation (Figure 4.9-10). The Xist^{APID} cell line generated and presented above will be a good starting point to investigate this interplay of pathways, for example, by comparing SmcHD1 dynamics in this cell line with the altered dynamics observed in the *Lrifi*^{-/-} cell line.

Another aspect to consider is that LRIF1 enrichment over Xi depends on SmcHD1 in cells which have undergone XCI (Brideau et al., 2015; Ichihara et al., 2022). It would be interesting to determine whether this is the case also during the establishment phase of XCI and is, thus, a general feature of the association between SmcHD1, LRIF1 and heterochromatic regions, as actually was observed for the pericentric regions and discussed above.

Chapter 5: Dissecting dynamic behaviour of SmcHD1 using SmcHD1 mutants

5.1 Introduction

Having set up a system which allows the study of SmcHD1 dynamic behaviour, including specifically across the Xi chromosome (Chapter 3), and having successfully used it to gain more information about the role of the SmcHD1-LRIF1-HP1-H3K9me3 pathway, both generally across the nucleus and in relation to XCI (Chapter 4), I went on to employ this methodology further to study several different SmcHD1 mutants, to understand better the relationship between SmcHD1 and underlying chromatin.

As described above (Section 1.3.1), SmcHD1 is a non-canonical member of the Smc family of chromosomal proteins (Blewitt et al., 2008). Canonical Smc proteins are motor proteins and rely on ATP hydrolysis to fulfil their key role in chromosomal organisation (Section 1.1.3). Using a mutation, E147A, which abolishes the GHKL ATPase activity of SmcHD1 revealed a requirement for enriched localization over the Xi (Brideau et al., 2015). However, exactly how this activity influences recruitment of SmcHD1 to its genomic targets and its functionality has not been investigated so far. For example, what would happen if the ATPase activity were increased rather than decreased? Thus, I wanted to focus my study on mutations affecting this reaction.

The first mutation I investigated was E147A, mentioned above, which neutralizes the ATPase activity, by switching residue 147 glutamic acid to an alanine (Brideau et al., 2015; K. Chen, R. C. Dobson, et al., 2016). This residue is highly conserved among GHKL ATPases, and it is located in the ATP-binding pocket, being one of the three catalytic residues (Panaretou et al., 1998; Pedersen et al., 2019; Prodromou et al., 1997). Importantly, it seems that the E147A mutation affects the hydrolysis step, rather than the ATP binding step of this process, since a crystal structure of human SMCHD1's GHKL ATPase domain containing this mutation has been solved in the presence of ATP (Pedersen et al., 2019).

A second mutation that I have investigated is the substitution of residue 137 glycine with glutamic acid (G137E) in the ATPase domain. This mutation is associated with BAMS, although it has been found also in one FSHD2 patient. *In vitro* analysis of a recombinant protein containing the SmcHD1 GHKL ATPase domain with this mutation showed an increase in its ATPase activity, both for a murine and a human construct (Gurzau et al., 2018; Pedersen et al., 2019), which is consistent with G137E being a BAMS-associated mutation. Notably, the mutation did not affect the protein's stability, assessed by a thermal denaturation assay, and injection in *Xenopus* embryos of mRNA encoding for human SMCHD1 having the G137E mutation resulted in a decreased eye diameter reminiscent of the microphthalmia phenotype seen in BAMS patients (Gurzau et al., 2018). G137 is part of a hot spot region, with residues 134-136 also undergoing gain-of-function mutations, and is located close to the catalytic motif I (that contains E147), oriented towards the outer interface of the protein (Gurzau et al., 2018). However, exactly how G137E influences the ATPase activity of SmcHD1 is not known.

Lastly, I set out to analyse a mutation found in FSHD2 patients which substitutes tyrosine 353 with cysteine (Y353C). This mutation leads to a dramatic decrease in the ATPase activity *in vitro* of a purified recombinant SmcHD1 GHKL ATPase domain, although

the thermal stability of the construct is not impaired compared to its WT counterpart (Gordon et al., 2017; Gurzau et al., 2018). The residue is oriented towards the outer surface of the protein and placed between the GHKL ATPase domain and its extended region. As for G137E, it is not clear why Y353C substitution impairs SmcHD1's ATPase activity.

One aspect to note is that all FSHD2- and BAMS-associated mutations are heterozygous in patients. In my ESC system, I will introduce them as homozygous to avoid confounding results due to potential formation of heterodimers between WT and mutant proteins and alleviation of some defects by the WT copies of the protein. This is especially important for looking at the effect of these mutations on SmcHD1's function in XCI since, in general, neither FSHD2 nor BAMS patients have been reported to display defects related to this process (Dion et al., 2019; Wang, Brand, et al., 2019).

5.2 Generation of a catalytically-dead, BAMS-associated and FSHD2-associated SmcHD1 mutant cell lines

For each of the three chosen mutations, E147A, G137E and Y353C, I used CRISPR/Cas9-mediated homologous recombination to integrate mutations at the endogenous *SmcHD1* locus in the WT cell line, T7Halo-SmcHD1 R26-eGFP-CIZ1 (Figure 5.1 A). The generation of the SmcHD1^{Y353C} mESC line was done together with the help of an undergraduate student whom I supervised, Victoria Pustygina. The obtained clones were initially screened by PCR and Sanger sequencing and, later on, the successful homozygous introduction of the desired mutation was confirmed by mapping the chrRNA-seq reads (Figures 5.1 B-D).

Western Blot from nuclear extracts was used to check whether the mutations have an impact on the protein's expression level. While G137E did not affect SmcHD1 levels (Figure

5.1 F), both E147A and Y353C caused a decrease in the amount of detectable protein (Figures 5.1 E, G). This reduction alone should not affect functionality of SmcHD1 since, as described in section 3.2, a C-terminal tagged SmcHD1, which has highly reduced expression levels compared to the untagged protein, displayed normal function in all assays performed. For each of the three cell lines, I selected for all further experiments one clone, marked with an asterisk in Figures 5.1 E-G, based on the protein's sequence analysis, expression level and quality of signal in live imaging.

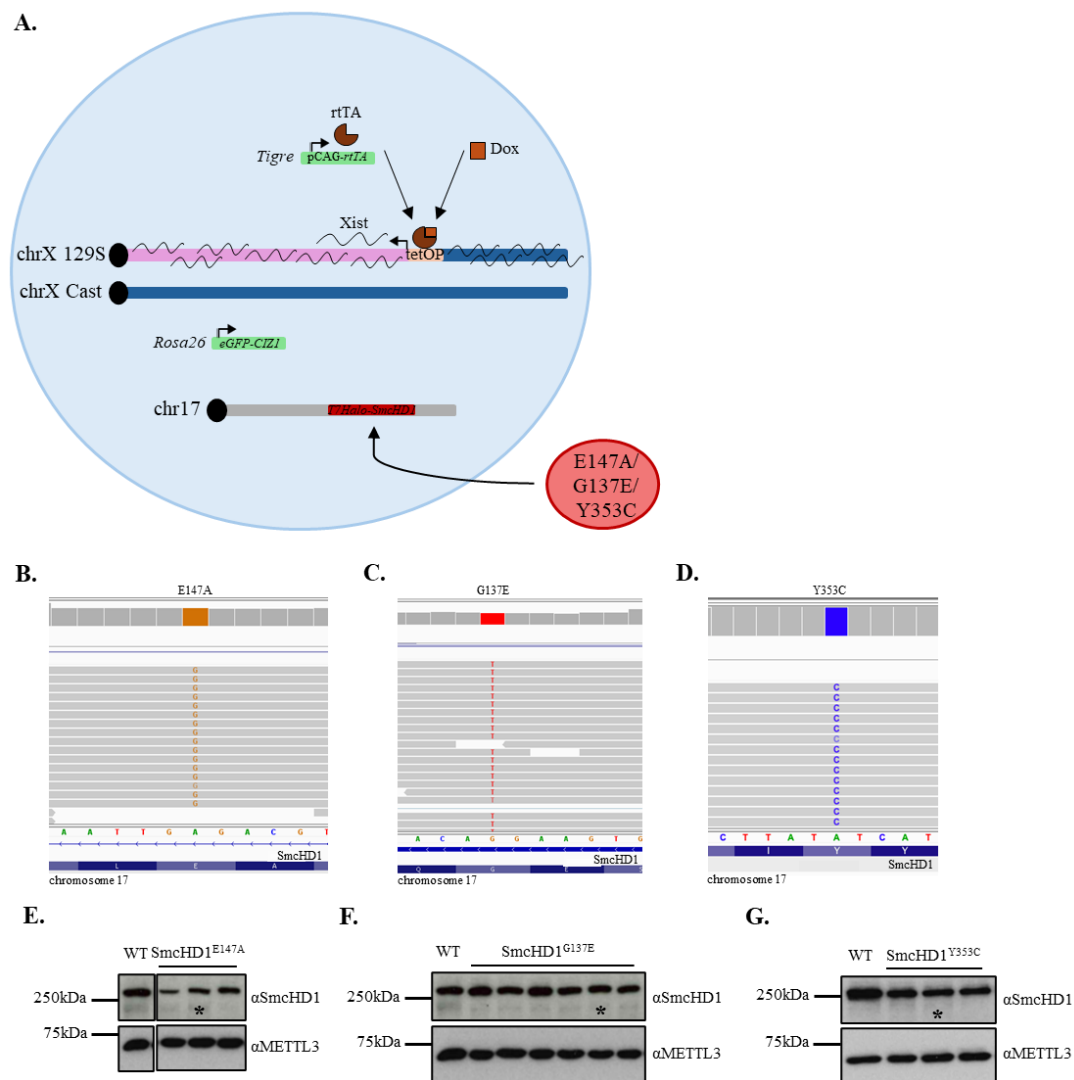


Figure 5.1 – Generation of SmcHD1 mutant mESC lines

Figure 5.1 – Generation of SmcHD1 mutant mESC lines:

A. Schematic showing the WT T7Halo-SmcHD1 R26-eGFP-CIZ1 mESC line and the generation of an endogenous SmcHD1^{E147A/G137E/Y353C} mutant mESC line. **B-D.** Genome browser screenshot of chrRNA-seq reads showing homozygosity of the chosen SmcHD1^{E147A} (**B**), SmcHD1^{G137E} (**C**) or SmcHD1^{Y353C} (**D**) clones. **E-G.** Western Blot showing the expression of SmcHD1 in nuclear extracts from WT and SmcHD1^{E147A} (**E**), SmcHD1^{G137E} (**F**) or SmcHD1^{Y353C} (**G**) clones. METTL3 acts as a loading control. * marks the clones which were selected for further experiments for each cell line. In (**E**), the extracts were run on the same gel but in non-adjacent lanes

5.3 Effects of SmcHD1 mutations on localization and dynamic behaviour over the whole nucleus

Having obtained the three SmcHD1 mutant cell lines, I was first interested in studying whether the general nuclear localization pattern of the protein is affected by these mutations.

Upon imaging the Halo-labelled SmcHD1^{E147A} cells, it could be seen that enrichment of SmcHD1 at pericentric regions is lost across the entire cell population (Figure 5.2 A). Analysis of the dynamic behaviour of the mutant protein across the whole nucleus revealed that SmcHD1^{E147A} has a larger bound fraction compared to the fast-moving fraction in both undifferentiated and differentiated cells (Figure 5.2 B-C). The proportion of bound molecules increases from ~42% for WT SmcHD1 to 51% at day 0 and to 61% at day 7 for SmcHD1^{E147A} (Figure 5.2 D-E). This result provides an interesting contrast with analysis of *Lrifi*^{-/-} cells where there is also a loss of pericentric localisation (Figure 4.2 A) but also complete abolition of the bound fraction (Figure 4.3).

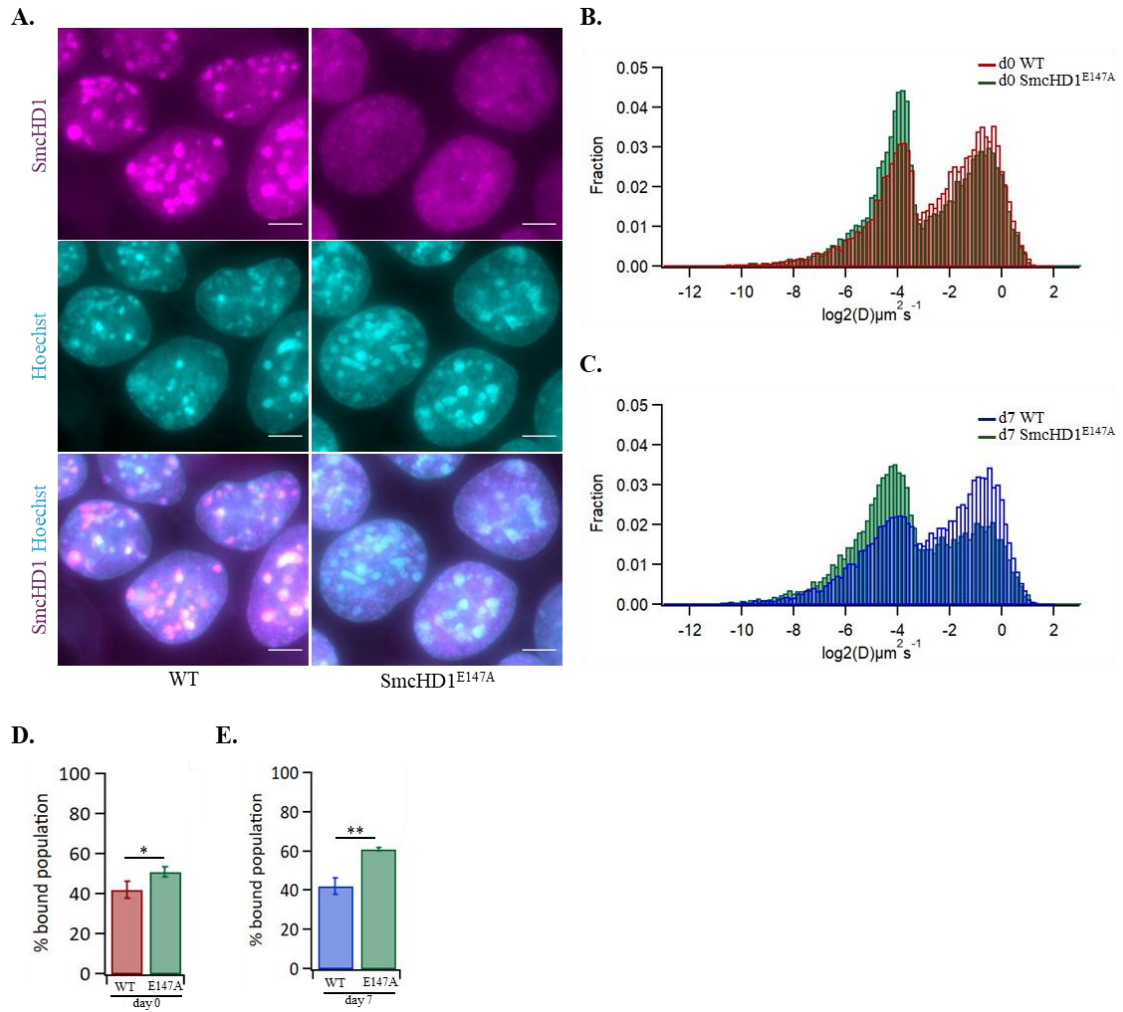


Figure 5.2 – The effect of SmcHD1^{E147A} mutation on the protein's localization to pericentric regions and dynamic behaviour across the whole nucleus:

A. Representative live images of Halo-tag labelling and Hoechst staining in WT and SmcHD1^{E147A} mESC lines, at day 0. Images are maximum-intensity projections, scale bar corresponds to 5μm. **B.** In green, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 11515 from 64 movies); the imaged cells are undifferentiated SmcHD1^{E147A} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **C.** In green, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 17509 from 66 movies); the imaged cells are SmcHD1^{E147A} mESCs which have undergone ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **D-E.** Bar plots showing the average percentage of bound molecules of SmcHD1 at day 0 (D) and day 7 differentiated (E) mESCs in WT and SmcHD1^{E147A} cells, from 3 independent replicates. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

In contrast to E147A, the G137E mutation did not impair enrichment of SmcHD1 at pericentric heterochromatin (Figure 5.3 A). Very interestingly though, analysis of SmcHD1^{G137E} dynamic behaviour revealed a skew towards the bound state in both analysed cellular conditions (Figure 5.3 B-C), similar to those observed for SmcHD1^{E147A} (Figure 5.2 B-C). Therefore, the proportion of bound molecules is larger for SmcHD1^{G137E}, at 57-59%, compared to WT SmcHD1 (Figure 5.3 D-E).

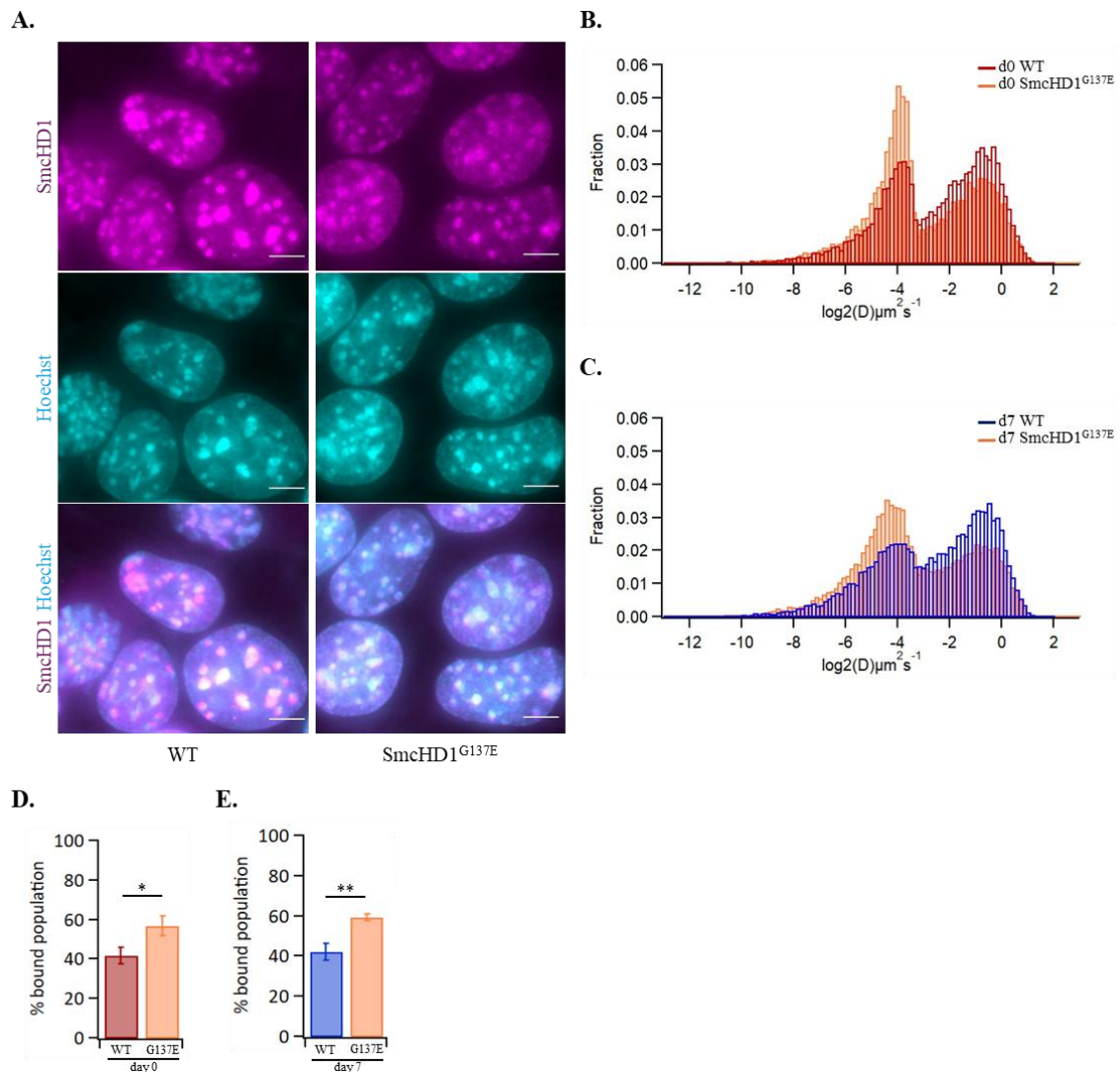


Figure 5.3 – The effect of the SmcHD1^{G137E} mutation on the protein's localization to pericentric regions and dynamic across the whole nucleus

Figure 5.3 – The effect of the SmcHD1^{G137E} mutation on the protein's localization to pericentric regions and dynamic across the whole nucleus:

A. Representative live images of Halo-tag labelling and Hoechst staining in WT and SmcHD1^{G137E} mESC lines, at day 0. Images are maximum-intensity projections, scale bar corresponds to 5µm. **B.** In orange, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 12374 from 64 movies); the imaged cells are undifferentiated SmcHD1^{G137E} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **C.** In orange, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 19309 from 66 movies); the imaged cells are SmcHD1^{G137E} mESCs which have undergone ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **D-E.** Bar plots showing the average percentage of bound molecules of SmcHD1 at day 0 (C) and day 7 differentiated (E) mESCs in WT and SmcHD1^{G137E} cells, from 3 independent replicates. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

Finally, imaging the SmcHD1^{Y353C} mESCs showed the Y353C mutation leads to loss of localized accumulation of SmcHD1 at the pericentric domains (Figure 5.4 A), similar to that seen in cells lacking LRIF1 (Figure 4.2 A) and cells with mutant SmcHD1^{E147A} (Figure 5.2 A). Regarding its dynamic behaviour across the nucleus, SPT revealed only modest changes compared to WT SmcHD1, with the mutant having slightly more fast-moving molecules, particularly at day 7 (Figure 5.4 B-C). Accordingly, the amount of bound molecules decreases to 35-39% (Figure 5.4 D-E). While seeing such small changes in the population distribution profile was surprising given the drastic change in SmcHD1 nuclear localization, it is important to note that this is only one aspect of a protein's dynamic behaviour and its residence time is also key for its mode of interaction with the underlying chromatin. This point will be addressed using FRAP in the following section.

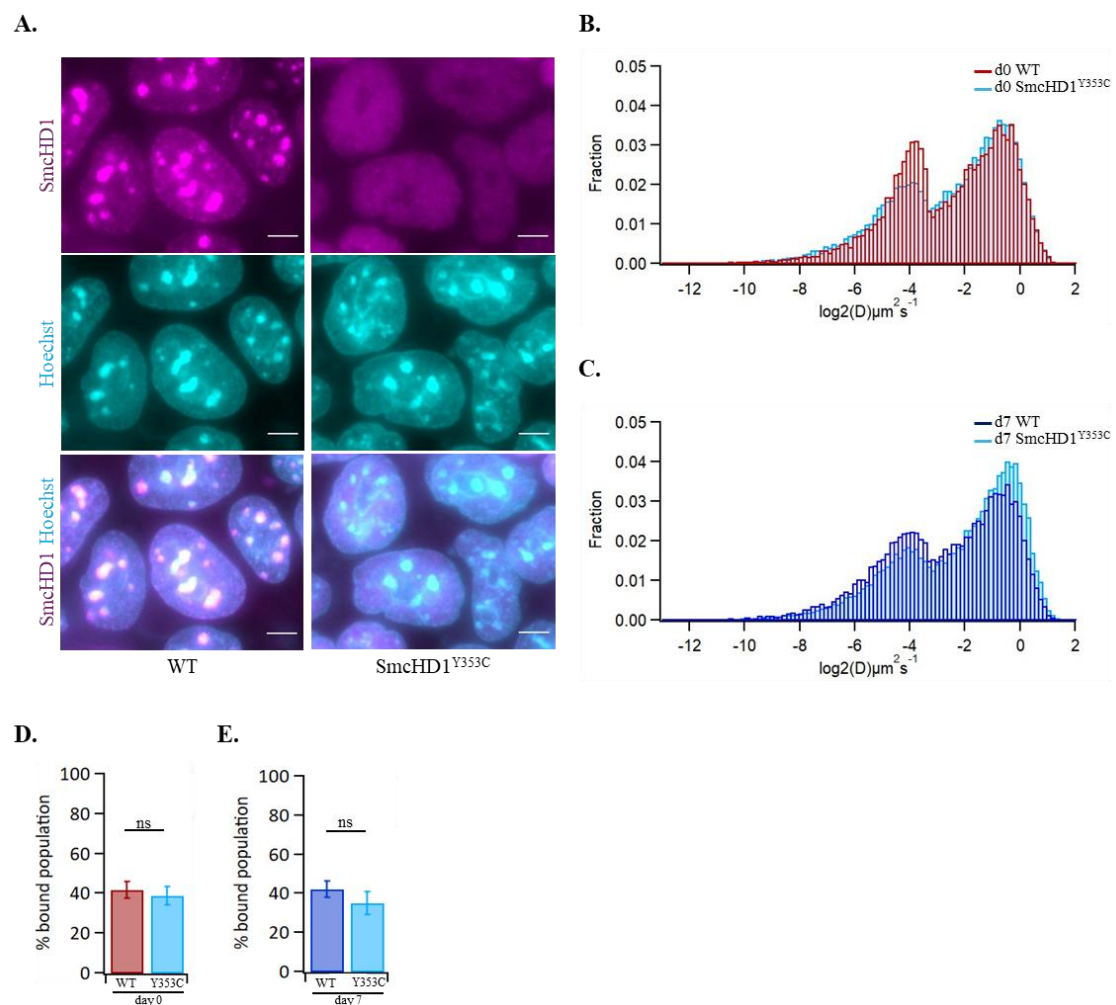


Figure 5.4 – The effect of SmcHD1^{Y353C} mutation on the protein’s localization to pericentric regions and dynamic behaviour across the whole nucleus:

A. Representative live images of Halo-tag labelling and Hoechst staining in WT and SmcHD1^{Y353C} mESC lines at day 0. Images are maximum-intensity projections, scale bar corresponds to 5μm. **B.** In turquoise, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 23820 from 66 movies); the imaged cells are undifferentiated SmcHD1^{Y353C} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **C.** In turquoise, histogram of diffusion coefficients of analysed tracks from across the whole nucleus, normalised for the number of analysed tracks (n = 32336 from 66 movies); the imaged cells are SmcHD1^{Y353C} mESCs which have undergone ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the whole nucleus, at the same timepoint. **D-E.** Bar plots showing the average percentage of bound molecules of SmcHD1 at day 0 (D) and day 7 differentiated (E) mESCs in WT and SmcHD1^{Y353C} cells, from 3 independent replicates. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

5.4 Effects of SmcHD1 mutations on recruitment and dynamic behaviour across the Xi chromosome

Having seen the nuclear wide effects that three SmcHD1 mutations have, I wanted then to focus specifically on the Xi domain.

Taking the SmcHD1^{E147A} mESCs through a 7 day ML differentiation live imaging timecourse, I observed the mutant has lost the ability to enrich over the Xi territory, although CIZ1 accumulation is retained (Figure 5.5 A). This is further confirmed by the quantification of the SmcHD1 signal intensity underneath the CIZ1 domain which shows lower values compared to WT at all analysed days (Figure 5.5 B). Of note is that there is a minority of cells at days 6-7 which show some accumulation of SmcHD1^{E147A} over the Xi region, although to a very low degree, similar to that observed for days 2-3 for the WT protein.

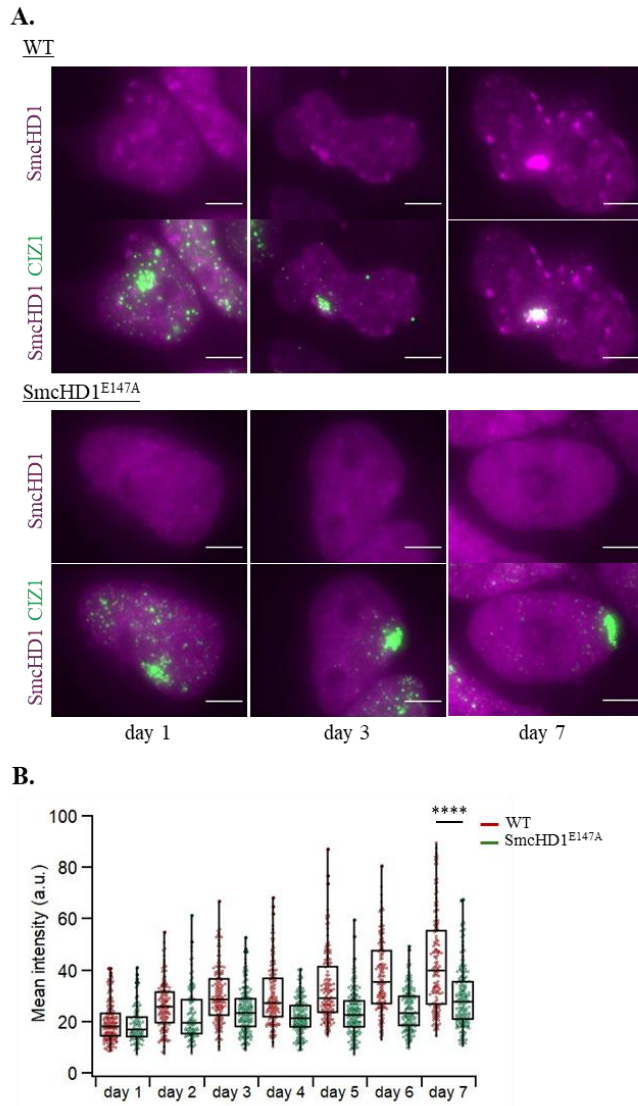


Figure 5.5 - The effect of SmcHD1^{E147A} mutation on SmcHD1 localization to the Xi chromosome:

A. Representative live images showing there is gradual enrichment or no enrichment of SmcHD1 over the Xi domain marked by eGFP-CIZ1 from day 1 to day 3 to day 7 of ML differentiation in WT and SmcHD1^{E147A} mESCs, respectively. Images are maximum-intensity projections, scale bar corresponds to 5μm. **B.** Quantification of SmcHD1 signal underneath the CIZ1 domain, at each day, during a 7 days live imaging ML differentiation timecourse comparing WT (red) and SmcHD1^{E147A} (green) mESCs. Each dot represents a cell. n = 102-200. Statistical analysis was done using the unpaired t-test.

Subsequently, looking at the dynamic behaviour of SmcHD1^{E147A} across the Xi chromosome, revealed a small decrease in the fast fraction compared to WT at day 0 (Figure 5.6 A), as was observed also across the whole nucleus (Figure 5.2 B). The fraction of bound

molecules increased, thus, from 51% for WT to 60% for SmcHD1^{E147A} (Figure 5.6 F). Importantly, the population distribution profile within the Xi is now very similar to that observed for a random nucleoplasmic region (Figure 5.6 B) and, therefore, the difference in dynamic behaviour that is usually observed between these regions in WT (Figure 3.4 E) is lost for the mutant SmcHD1^{E147A}. At day 7, across the Xi domain, there is almost no difference between WT and SmcHD1^{E147A} (Figure 5.6 C) and, since the bound population is already large at day 0, no change occurs from day 0 to 7 for the mutant (Figure 5.6 E). The amount of bound molecules for SmcHD1^{E147A} at day 7 is 63%, very close to the 60% observed for WT SmcHD1 (Figure 5.6 F). A random region in the differentiated SmcHD1^{E147A} cells shows an even distribution between the populations, with the fraction of bound molecules at 51%, and, thus, a bit lower than across the Xi region (Figure 5.6 D, F). That WT SmcHD1 and SmcHD1^{E147A} show similar fractions of bound molecules across the Xi seems counterintuitive given the absence of Xi SmcHD1 enrichment in SmcHD1^{E147A} cells (Figure 5.5), but the key observation is that the dynamic behaviour of SmcHD1 in the rest of the nucleoplasm in SmcHD1^{E147A} cells is similar to that seen over the Xi in WT cells (Figure 5.6 D, 3.4 F). The possible basis for this finding is discussed further below (section 5.6).

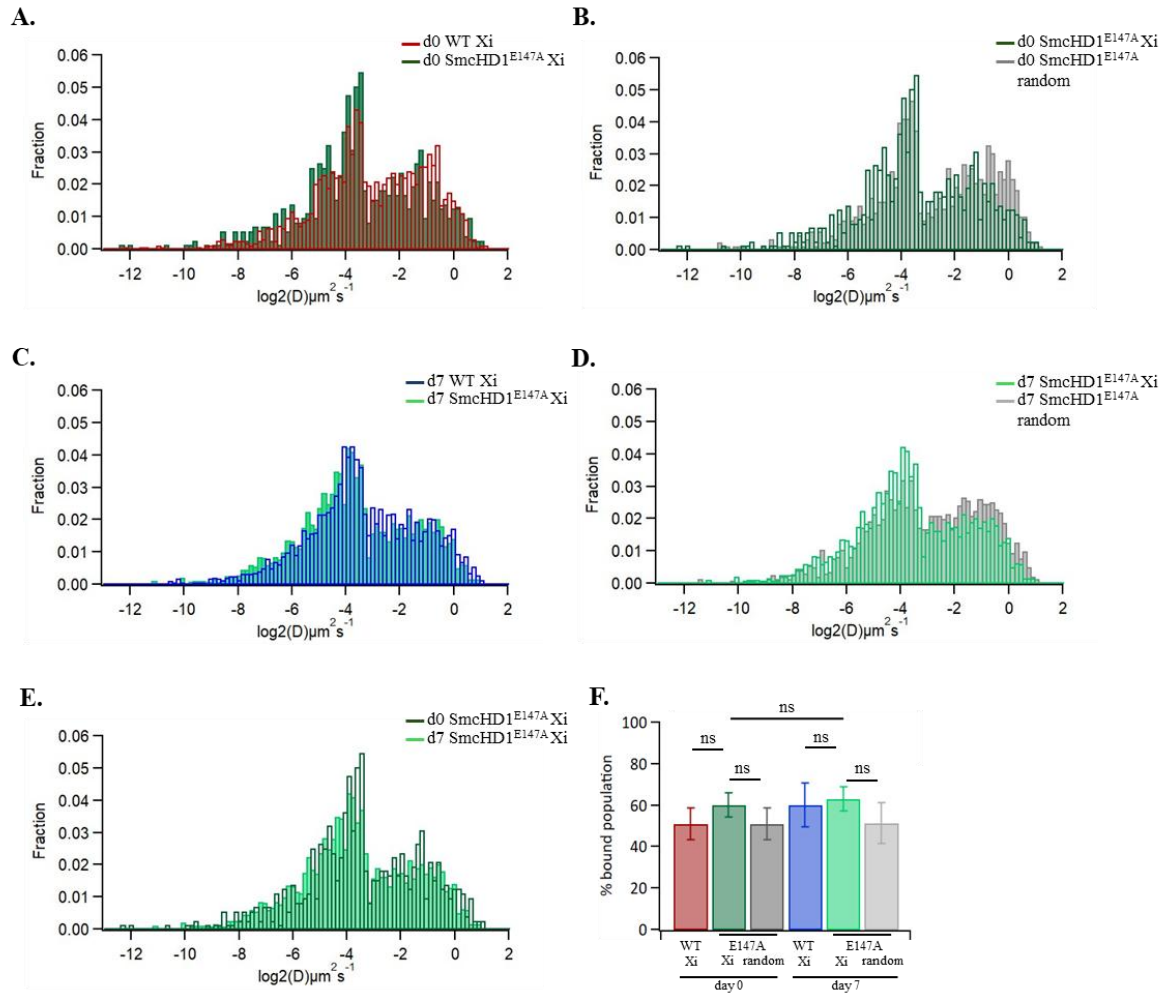


Figure 5.6 - The effect of SmcHD1^{E147A} mutation on the protein dynamic behaviour across the Xi chromosome:

A. In dark green, histogram of diffusion coefficients of analysed tracks from the Xi territory, normalised for the number of analysed tracks ($n = 712$ from 262 Xi); the imaged cells are undifferentiated SmcHD1^{E147A} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **B.** In dark grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 853$ from 262 random regions); the imaged cells are undifferentiated SmcHD1^{E147A} mESCs, kept for 24h with doxycycline. In dark green, same histogram as shown in A. **C.** In light green, histogram of diffusion coefficients of analysed tracks from across the Xi territory, normalised for the number of analysed tracks ($n = 2315$ from 242 Xi); the imaged cells are SmcHD1^{E147A} mESCs which underwent ML differentiation for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **D.** In light grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 1902$ from 242 random regions); the imaged cells are SmcHD1^{E147A} mESCs which underwent ML differentiation for 7 days. In light green, same histogram as shown in C. **E.** Comparison of the histograms corresponding to dynamic behaviour of SmcHD1 over the Xi territory at day 0 (dark green, from A) versus day 7 (light green, from C) in SmcHD1^{E147A} mESCs.

Figure 5.6 (continued) - The effect of SmcHD1^{E147A} mutation on the protein dynamic behaviour across the Xi chromosome:

F. Bar plot showing the average percentage of bound molecules of SmcHD1 at day 0 and day 7 differentiated mESCs over the Xi territory and a random nuclear region, in WT (5 independent replicates) and SmcHD1^{E147A} (4 independent replicates) mESCs. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

The G137E mutation does not impair enrichment of SmcHD1 over the Xi chromosome in differentiating cells (Figure 5.7 A). However, visual inspection of the cells, as well as quantification of the Xi-associated SmcHD1 intensity signal, indicate that initial Xi recruitment (days 1) is significantly accelerated (Figure 5.7 B).

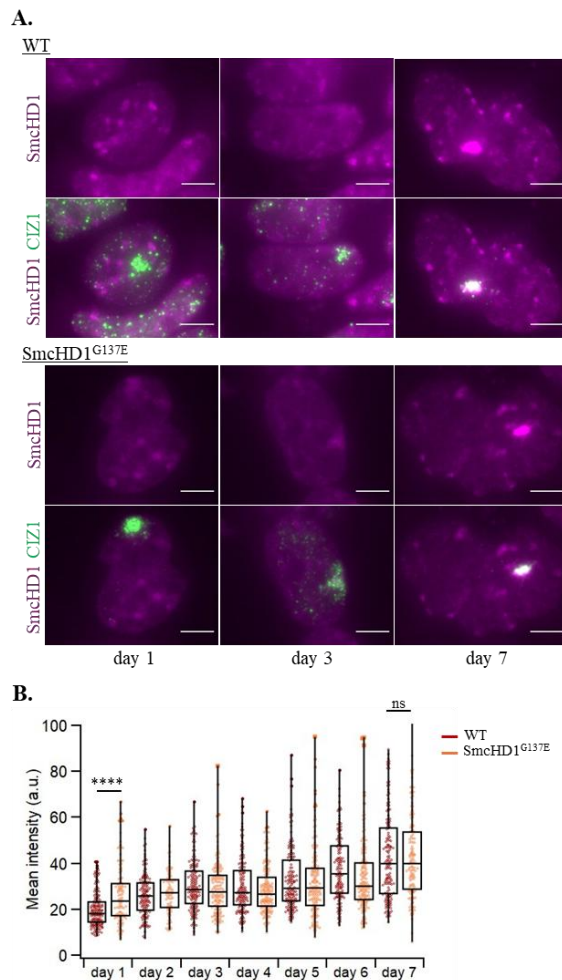


Figure 5.7 - The effect of SmcHD1^{G137E} mutation on Xi localization

Figure 5.7 - The effect of SmcHD1^{G137E} mutation on Xi localization:

A. Representative live images showing gradual accumulation of SmcHD1 over the Xi domain marked by eGFP-CIZ1 from day 1 to day 3 to day 7 of ML differentiation in WT and SmcHD1^{G137E} mESCs. Images are maximum-intensity projections, scale bar corresponds to 5µm. **B.** Quantification of SmcHD1 signal underneath the CIZ1 domain, at each day, during a 7 days live imaging ML differentiation timecourse comparing WT (red) and SmcHD1^{G137E} (orange) mESCs. Each dot represents a cell. n = 83-200. Statistical analysis was done using the unpaired t-test.

Regarding its dynamic behaviour measured by SPT, at day 0, SmcHD1^{G137E} presents a prominent bound fraction across the Xi domain, larger than that observed for WT SmcHD1 (Figure 5.8 A). In this condition, the proportion of bound molecules in the mutant is 69% (Figure 5.8 F). Comparing the Xi to a random nucleoplasmic region, SmcHD1^{G137E} retains a difference in dynamic behaviour between the two areas, with the latter having less measurable molecules in a bound state, at 52% (Figure 5.8 B, F). This fraction is much higher than reported for WT SmcHD1 (37%), but is in accordance with the area representing a subset of the whole nucleus where SmcHD1^{G137E} shows an increased bound population (Figure 5.3 B-C, 5.8 F). In day 7 differentiated SmcHD1^{G137E} cells, within the Xi territory, the mutant has the same population distribution profile as WT SmcHD1, with a 67% bound fraction (Figure 5.8 C, F). As in WT cells (Figure 3.4H), there is a clear difference between the Xi and a random region, with the latter having a significantly smaller bound population, at 41% (Figure 5.8 D, F). From day 0 to day 7, SmcHD1^{G137E} presents almost no change in its dynamic profile across the Xi domain (Figure 5.8 E).

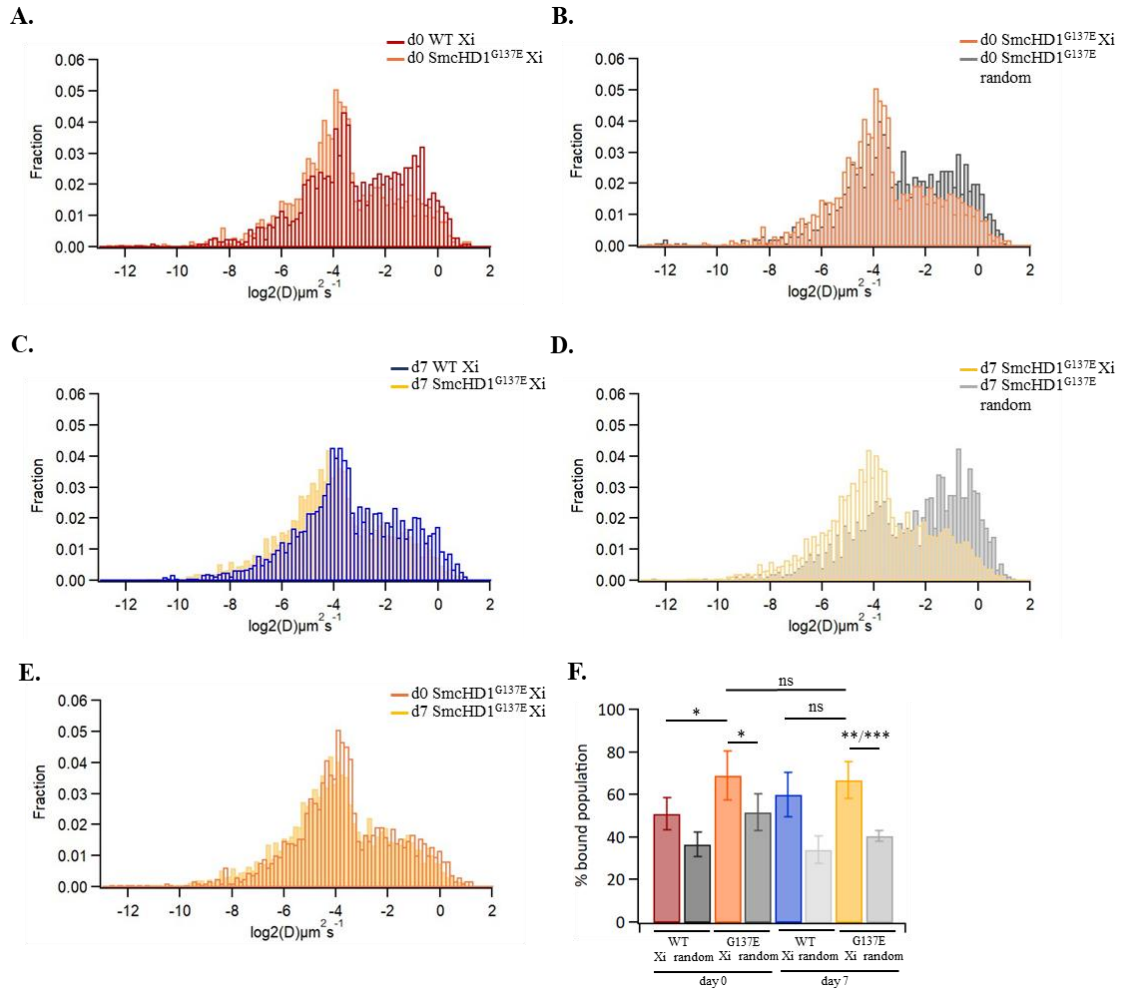


Figure 5.8 - The effect of SmcHD1^{G137E} mutation on the SmcHD1 dynamic behaviour across the Xi chromosome:

A. In orange, histogram of diffusion coefficients of analysed tracks from the Xi territory, normalised for the number of analysed tracks ($n = 1542$ from 262 Xi); the imaged cells are undifferentiated SmcHD1^{G137E} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **B.** In dark grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 945$ from 262 random regions); the imaged cells are undifferentiated SmcHD1^{G137E} mESCs, kept for 24h with doxycycline. In orange, same histogram as shown in A. **C.** In yellow, histogram of diffusion coefficients of analysed tracks from across the Xi territory, normalised for the number of analysed tracks ($n = 3896$ from 242 Xi); the imaged cells are SmcHD1^{G137E} mESCs ML differentiated for 7 days. In blue, histogram of diffusion coefficients showing dynamic behaviour of SmcHD1 in WT mESCs across the Xi territory, at the same timepoint. **D.** In light grey, histogram of diffusion coefficients of analysed tracks from across a random region in the nucleus, normalised for the number of analysed tracks ($n = 1405$ from 242 random regions); the imaged cells are SmcHD1^{G137E} mESCs ML differentiated for 7 days. In yellow, same histogram as shown in C. **E.** Comparison of the histograms corresponding to dynamic behaviour of SmcHD1 over the Xi territory at day 0 (orange, from A) versus day 7 (yellow, from C) in SmcHD1^{G137E} mESCs.

Figure 5.8 (continued) - The effect of SmcHD1^{G137E} mutation on the SmcHD1 dynamic behaviour across the Xi chromosome:

F. Bar plot showing the average percentage of bound molecules of SmcHD1 at day 0 and day 7 differentiated mESCs over the Xi territory and a random nuclear region, in WT (5 independent replicates) and SmcHD1^{G137E} (4 independent replicates) mESCs. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

Looking at the third mutant cell line, SmcHD1^{Y353C} mESCs, taken through ML differentiation for 7 days, it can be appreciated, from the images and associated quantification, that the mutation prevents the characteristic enrichment of SmcHD1 over the Xi domain with XCI progression (Figure 5.9). As observed for the *Lrifl*^{-/-} cells (Figure 4.4), there is an apparent increase in SmcHD1 signal intensity throughout the timecourse (Figure 5.9 B) which is not caused by accumulation of SmcHD1 across the Xi, but rather by a general increase in the signal across the entire nucleus (Figure 5.9 A). As previously discussed, further experiments would be required to confirm whether such changes in nuclear levels of SmcHD1 really take place due to the introduced mutation. One more aspect to note is that, as for SmcHD1^{E147A}, there is a small proportion of cells, at days 6-7, which show a modest level of SmcHD1 accumulation over the Xi territory, similar to what is observed at days 2-3 for the WT protein.

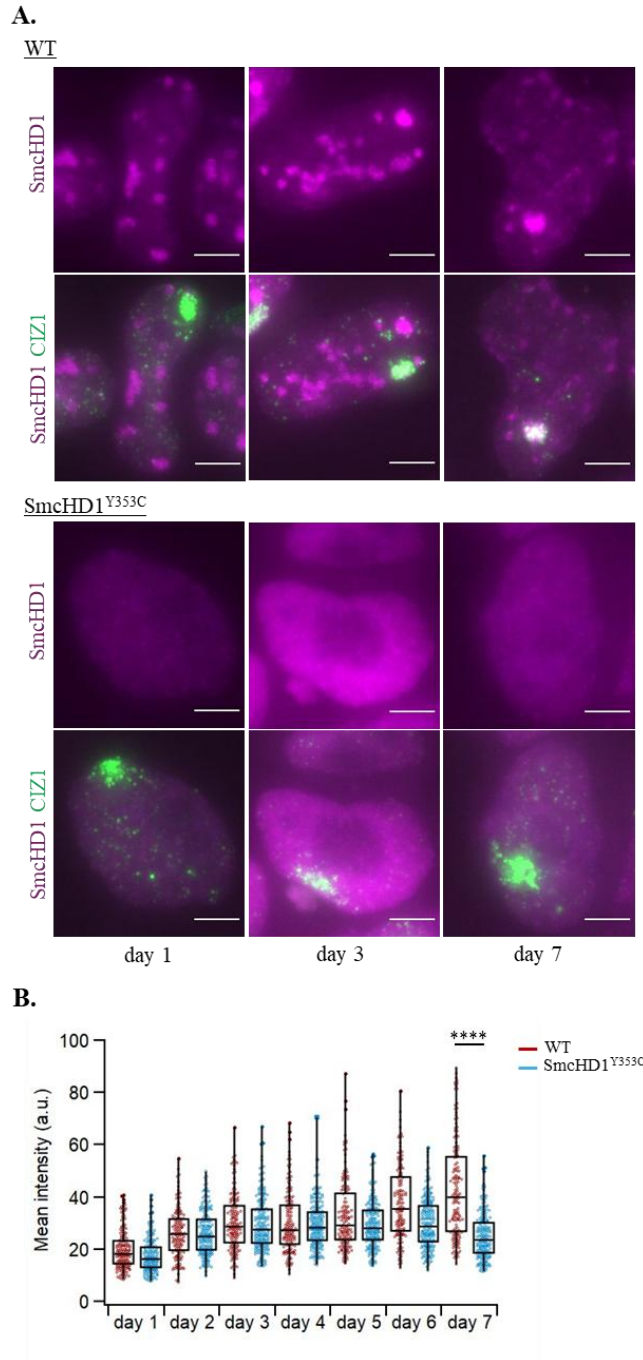


Figure 5.9 - The effect of SmcHD1^{Y353C} mutation on SmcHD1 localization to the Xi chromosome:

A. Representative live images showing there is gradual enrichment or no enrichment of SmcHD1 over the Xi domain marked by eGFP-CIZ1 from day 1 to day 3 to day 7 of ML differentiation in WT and SmcHD1^{Y353C} mESCs, respectively. Images are maximum-intensity projections, scale bar corresponds to 5µm. **B.** Quantification of SmcHD1 signal underneath the CIZ1 domain, at each day, during a 7 days live imaging ML differentiation timecourse comparing WT (red) and SmcHD1^{Y353C} (blue) mESCs. Each dot represents a cell. n = 136-206. Statistical analysis was done using the unpaired t-test.

Dynamic behaviour of SmcHD1^{Y353C} at day 0 in the context of the Xi chromosome were not affected compared to WT, with the amount of bound molecules being 55% (Figure 5.10 A, F). Consequently, the population distribution profile remains strikingly different between the Xi and a random region, with the latter having a much smaller bound population, at 35% (Figure 5.10 B, F). However, at day 7, within the Xi domain, the bound state is no longer prominent, as was seen for WT SmcHD1, with the majority of the SmcHD1^{Y353C} molecules being in the fast-moving state (Figure 5.10 C). The fraction of bound molecules is reduced to 48%, which is smaller compared to day 0, and, thus, the dynamic profile of the mutant across the Xi changes from day 0 to day 7 in the opposite direction to that seen for WT SmcHD1 (Figure 5.10 E-F, 3.4 J). Furthermore, at day 7, comparing the profile of SmcHD1^{Y353C} over the Xi territory with a random region, the difference in dynamic behaviour usually observed in WT (Figure 3.4 H) is greatly diminished, with both regions now displaying a higher peak for the fast than for the bound state (Figure 5.10 D, F).

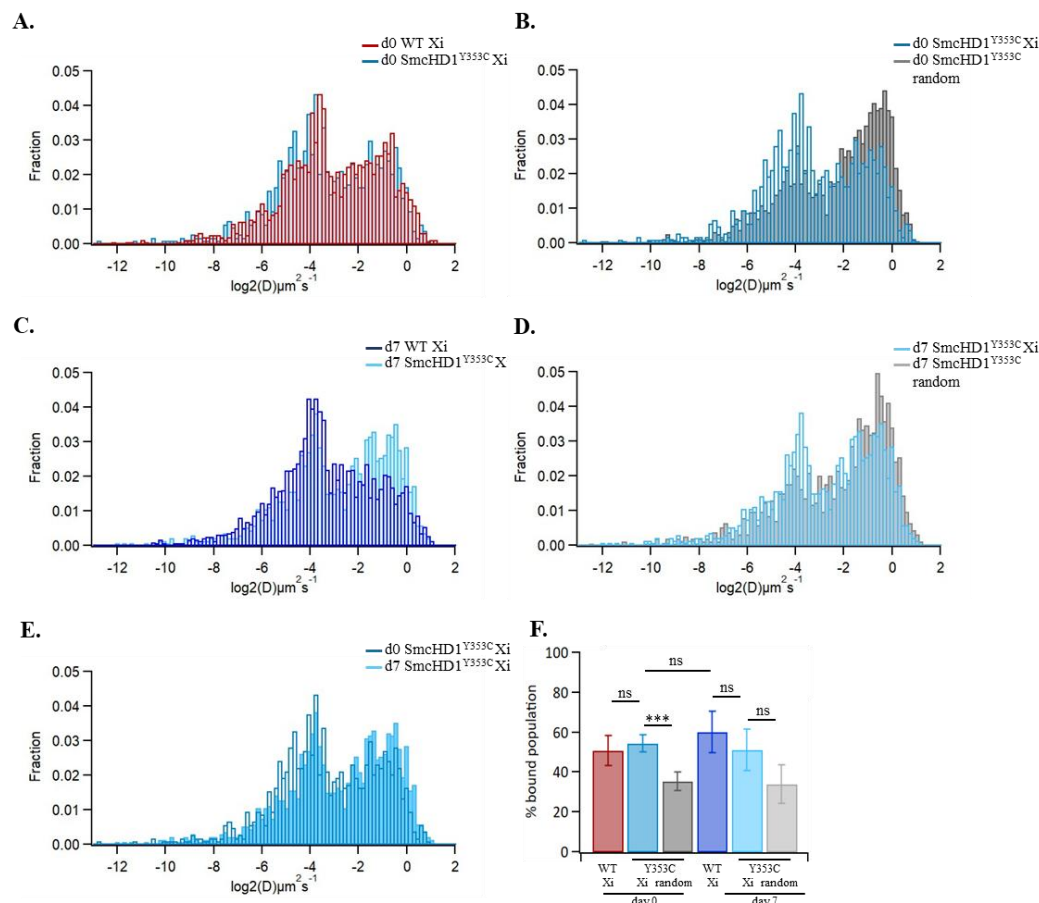


Figure 5.10 - The effect of *SmcHD1*^{Y353C} mutation on *SmcHD1* dynamic behaviour across the Xi chromosome:

A. In blue, histogram of diffusion coefficients of analysed tracks from the Xi territory, normalised for the number of analysed tracks ($n = 1033$ from 246 Xi); the imaged cells are undifferentiated *SmcHD1*^{Y353C} mESCs, kept for 24h with doxycycline. In red, histogram of diffusion coefficients showing dynamic behaviour of *SmcHD1* in WT mESCs across the Xi territory, at the same timepoint. **B.** In dark grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 1555$ from 246 random regions); the imaged cells are undifferentiated *SmcHD1*^{Y353C} mESCs, kept for 24h with doxycycline. In blue, same histogram as shown in A. **C.** In light blue, histogram of diffusion coefficients of analysed tracks from across the Xi territory, normalised for the number of analysed tracks ($n = 1331$ from 195 Xi); the imaged cells are *SmcHD1*^{Y353C} mESCs ML differentiated for 7 days. In dark blue, histogram of diffusion coefficients showing dynamic behaviour of *SmcHD1* in WT mESCs across the Xi territory, at the same timepoint. **D.** In light grey, histogram of diffusion coefficients of analysed tracks from a random region in the nucleus, normalised for the number of analysed tracks ($n = 1527$ from 195 random regions); the imaged cells are *SmcHD1*^{Y353C} mESCs ML differentiated for 7 days. In light blue, same histogram as shown in C. **E.** Comparison of the fits corresponding to dynamic behaviour of *SmcHD1* over the Xi territory at day 0 (blue, from A) versus day 7 (light blue, from C) in *SmcHD1*^{Y353C} mESCs. **F.** Bar plot showing the average percentage of bound molecules of *SmcHD1* at day 0 and day 7 differentiated mESCs over the Xi territory and a random nuclear region, in WT (5 independent replicates) and *SmcHD1*^{Y353C} (4 independent replicates) mESCs. Error bars represent the STDEV. Statistical analysis was done using the unpaired t-test.

To complement the SPT datasets and gain information about the residence time on chromatin of the three mutants, I also performed FRAP experiments. Snapshots of a FRAP experiment in WT cells and a detailed comment on the associated average fluorescence recovery curves is reported in section 4.3. In SmcHD1^{E147A} mESCs, it can be seen from both the snapshots shown in Figure 5.11 A and the plotted recovery curves, that the mutant recovers in a similar manner in the Xi domain and in the nucleoplasm (Figure 5.11 D-E). Therefore, the Xi-specific kinetic behaviour observed for WT SmcHD1 is lost (Figure 4.6 A, 5.11 D-E), as suggested also by the SPT study. In both areas, the recovery reaches about 40-50% of the pre-bleach fluorescence levels, slightly higher than the recovery level of WT SmcHD1 in the Xi territory. The turnover is on a timescale of seconds as, within the first minute, the fluorescence level goes up close to the final recovered level (Figure 5.11 D-E).

As in WT SmcHD1, SmcHD1^{G137E} recovers to a lower extent in the Xi domain compared to nucleoplasm (Figure 5.11 B). However, in both regions, the recovery level is much lower than their WT counterpart, reaching barely 10% of the pre-bleach fluorescence level in the Xi and 55% in the nucleoplasm (Figure 5.11 D-E). In accordance with the SPT data, this suggests the bound population of this mutant is in general higher. The slope of the average fluorescence recovery curves shows the rate of recovery is slower in both areas compared to that of WT SmcHD1: in about 3 minutes, the WT protein gets to around 15% pre-bleach fluorescence level over the Xi and 80% in the nucleoplasm, while, in the same time, SmcHD1^{G137E} only achieves about 6% recovery in the Xi and 40% in the nucleoplasm (Figure 5.11 D-E).

SmcHD1^{Y353C} presents fast dynamics, with the fluorescence recovering to almost pre-bleach levels, over both the Xi and nucleoplasm (Figure 5.11 C). As for SmcHD1^{E147A} and as shown via SPT, the difference in dynamics between the two regions, characteristic of WT SmcHD1, is no longer present, as can be seen in the average fluorescence recovery

curves shown in Figure 5.11 D-E. The recovery reaches 80-90% of the pre-bleach fluorescence level and this happens within the first minute after the bleach pulse, suggesting that the turnover is faster than that of WT SmcHD1, being on a timescale of seconds, rather than minutes (Figure 5.11 D-E).

As described in section 4.3, I chose to fit the average fluorescence recovery curves with a single exponential function to extract information about the immobile fraction and lifetime on chromatin of each analysed condition (Figure 5.11 E-H). Choosing the single over the biexponential fit was determined also by the very slow turnover of the SmcHD1^{G137E} mutant and the very fast turnover of SmcHD1^{Y353C} for which fitting with the latter function was not appropriate. It is also important to mention that the fitting for the curve corresponding to recovery of SmcHD1^{E147A} over the nucleoplasm is not accurate regarding its slope, with the lifetime being considerably overestimated, although this fit is better than a biexponential fit.

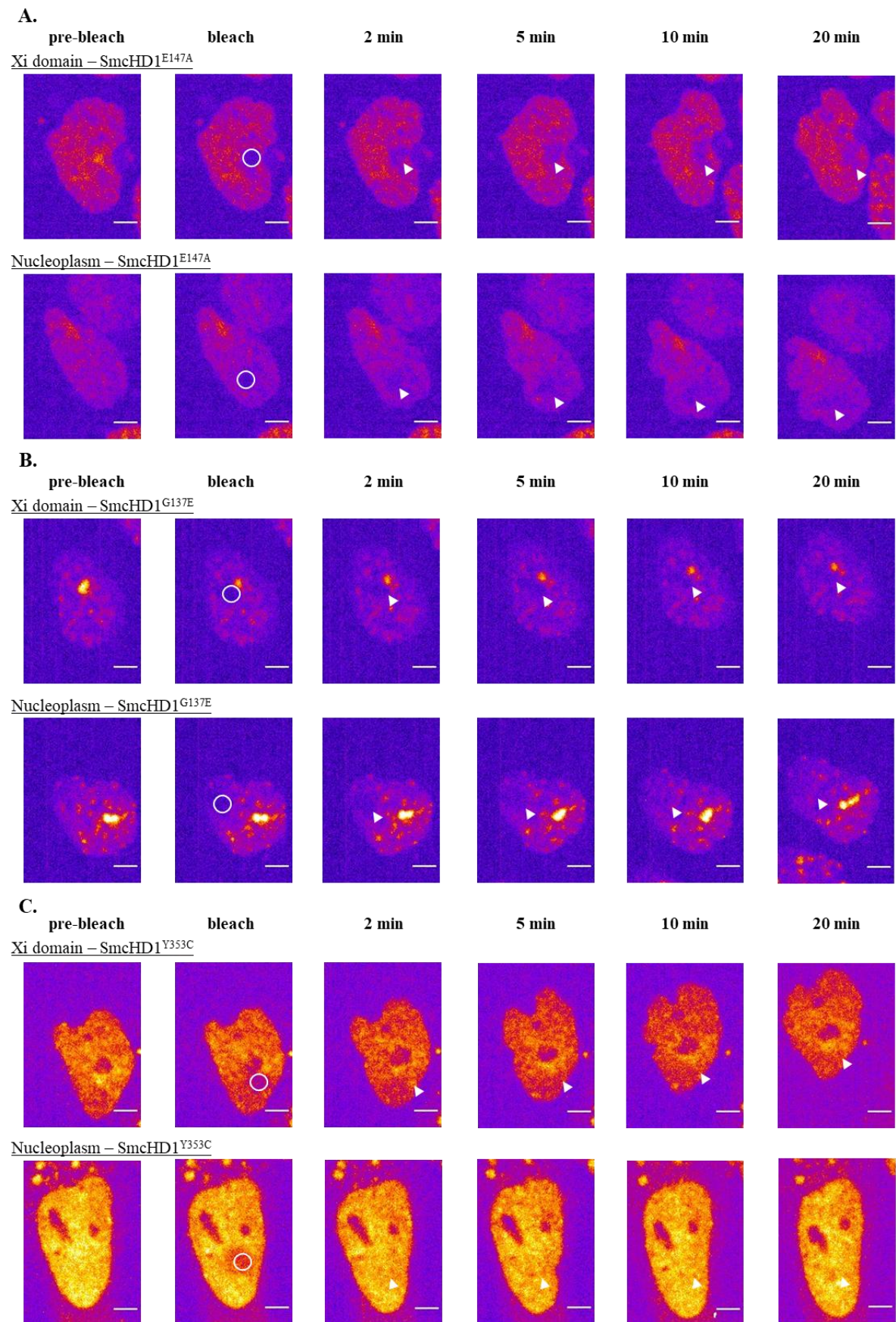


Figure 5.11 – The effect of SmcHD1 mutations on dynamic behaviour assessed by FRAP

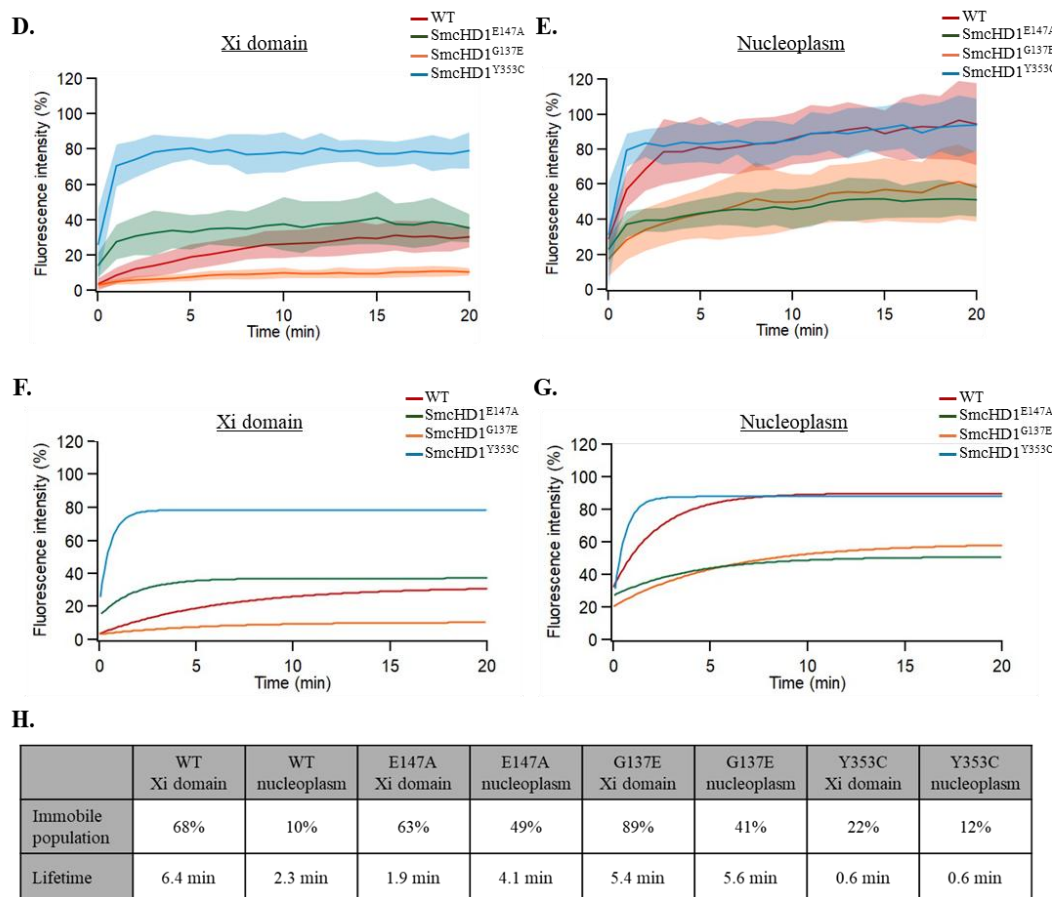


Figure 5.11 (continued) – The effect of SmcHD1 mutations on dynamic behaviour assessed by FRAP:

A-C. Representative images from a FRAP experiment using the SmcHD1^{E147A} (**A**), SmcHD1^{G137E} (**B**) or SmcHD1^{Y353C} (**C**) mESC which underwent ML differentiation for 7 days. The bleached region is marked by a white circle on the bleach snapshot and by a white arrow in the post-bleach snapshots. Scale bar corresponds to 5µm. **D-E.** Average fluorescence recovery curves over the Xi domain (**C**) or a random nucleoplasmic region (**D**) comparing WT, SmcHD1^{E147A}, SmcHD1^{G137E} and SmcHD1^{Y353C} mESCs which underwent ML differentiation for 7 days. The shading represents the STDEV. The number of cells analysed are Xi domain WT (16), SmcHD1^{E147A} (15), SmcHD1^{G137E} (15), SmcHD1^{Y353C} (14) and nucleoplasm WT (15), SmcHD1^{E147A} (23), SmcHD1^{G137E} (17), SmcHD1^{Y353C} (13). **F-G.** Single exponential fits of the average fluorescence recovery curves over the Xi domain (**F**) or a random nucleoplasmic region (**G**). **H.** Summary of the immobile population and lifetime values extracted from the fits shown in F-G.

The fits show that, while the E147A mutation does not affect the immobile fraction over the Xi, it crucially causes a decrease in the lifetime by about 3-fold (Figure 5.11 H). The Y353C mutation determines a reduction in both the immobile population and an even

more dramatic one in its residency time across the Xi chromosome (Figure 5.11 H). Lastly, the immobile fraction is greatly increased in both analysed areas in cells having the G137E mutation and this is accompanied by a ~2.5-fold increase in the protein's lifetime in the nucleoplasm. Such an increase probably takes place also within the Xi territory, but the extremely low recovery level prevents a good estimation of this time (Figure 5.11 F, H).

5.5 Effects of SmcHD1 mutations on its functionality during XCI

In parallel with analysing the impact of the three mutations on dynamic behaviour of SmcHD1, I wanted to assess whether the functions of SmcHD1 are altered.

Firstly, using chrRNA-seq, I checked for any effect on SmcHD1's role in X-linked gene silencing. Analysing all the allelically assignable genes across the X chromosome, all the cell lines showed the expected allelic ratio of ~0.5 in uninduced conditions and ~0.25 following Xist expression for 24h (Figure 5.12 A). However, at day 7, the allelic ratio for the SmcHD1^{G137E} cells is close to zero, as for WT SmcHD1, while cells having the E147A or Y353C mutation display a slightly higher allelic ratio (Figure 5.12 A). Splitting the genes into the three SmcHD1-dependency categories (Bowness et al., 2022), it can be appreciated that the SmcHD1-dependent genes are the most derepressed, with the SmcHD1-partially dependent genes being also affected, albeit less (Figure 5.12 B). For the SmcHD1^{G137E} mutant, there is no difference in the silencing of these two categories of genes. Moreover, for this mutant, given the observed accelerated accumulation over the Xi domain (Figure 5.7), I wanted to check whether this translates into accelerated gene silencing. At day 5, a modestly lower allelic ratio can be seen for SmcHD1^{G137E} cells compared to WT (Figure 5.12 A), with mainly the SmcHD1-dependent genes being more silenced (Figure 5.12 C).

Given the very small effect and the fact that this experiment was only carried out once, a further replicate would be needed for confirmation.

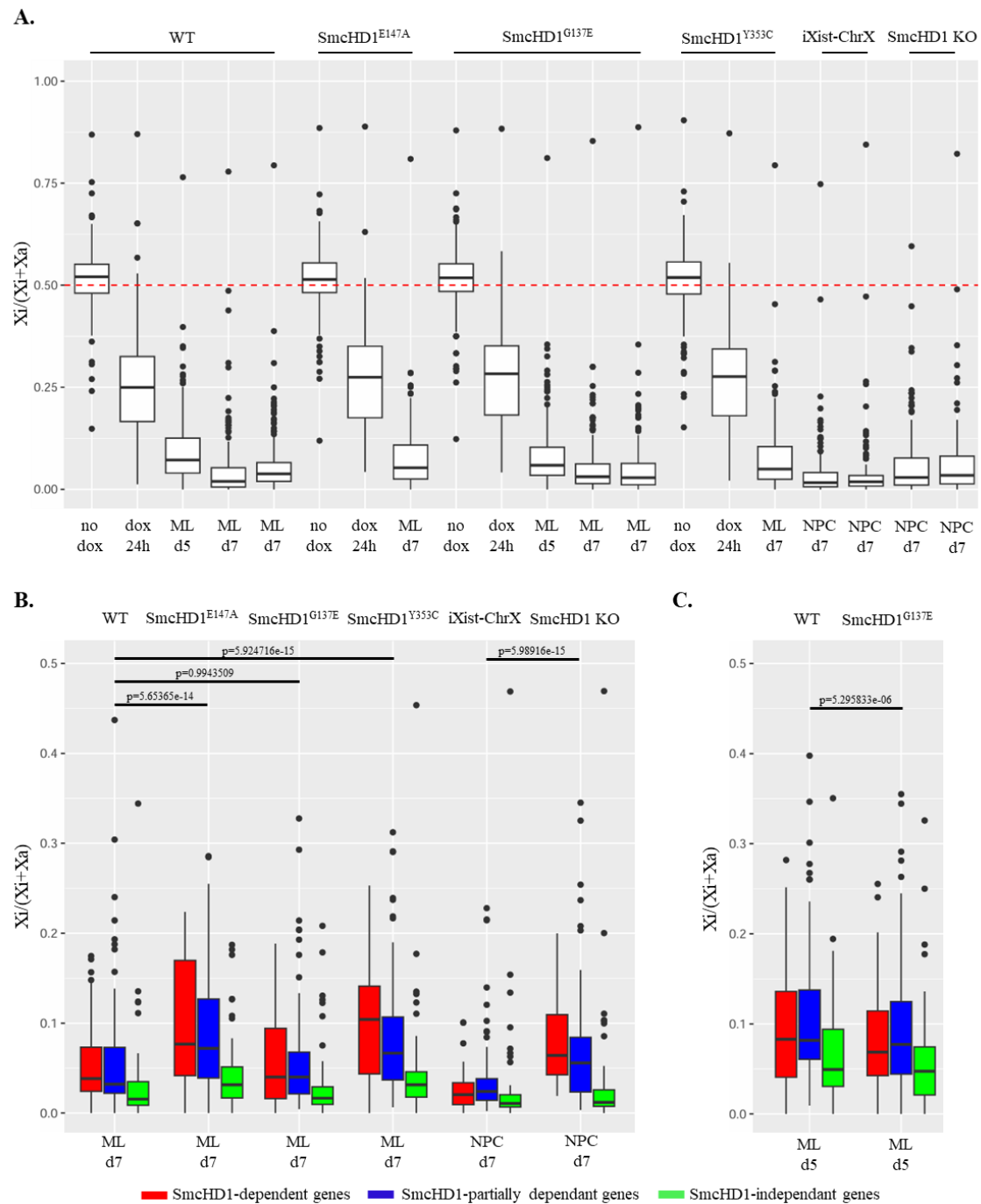


Figure 5.12 – The effect of SmcHD1 mutations on X-linked gene silencing assessed by chrRNA-seq

Figure 5.12 – The effect of SmcHD1 mutations on X-linked gene silencing assessed by chrRNA-seq:

A. Box plot showing allelic chrRNA-seq analysis of transcription of X-linked genes in WT, SmcHD1^{E147A}, SmcHD1^{G137E}, SmcHD1^{Y353C}, iXist-ChrX and SmcHD1 KO mESCs which were uninduced, induced for 24h with doxycycline, ML differentiated for 5 or 7 days or NPC differentiated for 7 days. iXist-ChrX is the parental cell line of the SmcHD1 KO cell line. **B-C.** Box plot showing allelic chrRNA-seq analysis of transcription of X-linked genes in WT, SmcHD1^{E147A}, SmcHD1^{G137E}, SmcHD1^{Y353C}, iXist-ChrX and SmcHD1 KO mESCs which were ML differentiated for 5 or 7 days or NPC differentiated for 7 days, with genes split in three categories according to their dependency on SmcHD1 for silencing. The WT ML d7, G137E ML d7, iXist-ChrX NPC d7 and SmcHD1 KO NPC d7 samples represent an average of two replicates, while the other samples represent one replicate. Statistical analysis was done using the paired t-test.

Secondly, I was interested in assessing whether the volume of the Xi-associated H3K27me3 domain is affected. In this respect, I carried out immunostaining with an antibody for H3K27me3 in cells which had undergone 7 days of ML differentiation (Figure 5.13 A). H3K27me3 is enriched over the Xi territory from the early stages of XCI (Plath et al., 2003; Silva et al., 2003). Quantification of the H3K27me3 domains showed a significant change only in the SmcHD1^{G137E} cells, where the volume of the domain is smaller (Figure 5.13 B). The other two mutants did not impact the Xi-associated H3K27me3 volume, if anything, the SmcHD1^{Y353C} cells had slightly larger domains (Figure 5.13 B).

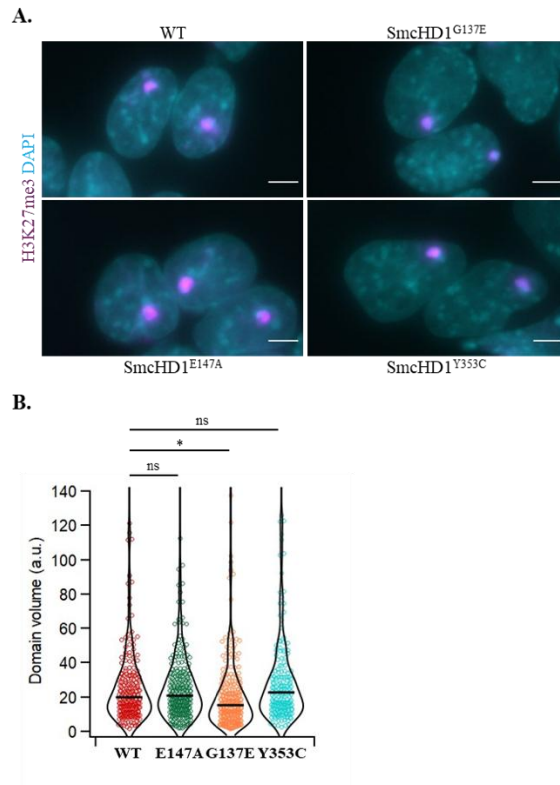


Figure 5.13 – The effect of SmcHD1 mutations on Xi-associated H3K27me3 domain volume:

A. Immunostaining for H3K27me3 in WT, SmcHD1^{E147A}, SmcHD1^{G137E} and SmcHD1^{Y353C} mESCs which have undergone ML differentiation for 7 days. Images are maximum-intensity projections, scale bar corresponds to 5µm. **B.** Violin plot showing the volume of the H3K27me3 domain in WT (red), SmcHD1^{E147A} (green), SmcHD1^{G137E} (orange) and SmcHD1^{Y353C} (blue) mESCs ML differentiated for 7 days. Each dot represents a cell. n = 209 (WT), 239 (SmcHD1^{E147A}), 259 (SmcHD1^{G137E}), 178 (SmcHD1^{Y353C}). Statistical analysis was done using the unpaired t-test.

5.6 Discussion

In this chapter, I assessed the impact of three mutations affecting SmcHD1's ATPase activity on the protein's localization and dynamic behaviour, both across the entire nucleus and specifically over the Xi chromosome. I further paired this information with an enquiry into

whether these changes lead to altered functionality of SmcHD1, particularly regarding its role in XCI.

For this study, I generated three mESC lines, each one having one of three chosen mutations: the loss-of-function mutations SmcHD1^{E147A} and SmcHD1^{Y353C} and the gain-of-function mutation SmcHD1^{G137E} (Figure 5.1). Importantly, the amino acid substitutions were homozygously introduced at the endogenous locus of *SmcHD1*, preventing side effects that may arise from a transgenic expression, in which case the level of expression as well as the mechanism of regulation may differ from that of the native locus. Any changes in the protein level of SmcHD1 are a direct consequence of the introduced mutation and consistent across all analysed clones. In this respect, it is interesting to note that the two mutants, SmcHD1^{E147A} and SmcHD1^{Y353C}, that destabilize SmcHD1's interaction with chromatin by decreasing its residence time are also the two that cause a reduction in the protein levels (Figure 5.1, 5.11). This is similar to the outcome in *Lrifi*^{-/-} cells, although there the level of SmcHD1 is even more strongly reduced, probably due to effects on both the bound population and the lifetime on chromatin being severely decreased (Figure 4.2, 4.3, 4.6).

Importantly, both SmcHD1^{E147A} and SmcHD1^{Y353C} mutants displayed the same phenotypic defects in relation to the nuclear localization of SmcHD1, causing loss of enrichment at the heterochromatic pericentric regions (Figure 5.2 A, 5.4 A) and over the Xi domain (Figure 5.5, 5.9). However, in each case, the altered pattern is linked to a different dynamic behaviour. This highlights the importance of dissecting different parameters underlying the kinetics of a protein of interest as performed in this study.

The E147A mutation causes a nuclear-wide increase in the bound population of SmcHD1 as assessed by SPT (Figure 5.2 B-E) and FRAP over random nucleoplasmic regions (Figure 5.11). At first glance, this may seem counterintuitive given the loss of

pericentric accumulation, but, looking at the specific dynamics within the Xi territory, allows us to observe the variation in dynamic behaviours within different nuclear regions and understand better the observed increase. The amount of SmcHD1^{E147A} bound to the Xi is not impaired at all at day 7 and is actually increased at day 0, to the same level seen for WT SmcHD1 at day 7 (Figure 5.6). However, what is important to note is that the bound population also goes up, to the same level, in a random nuclear region (Figure 5.6, 5.11), where the chances of including only genomic targets (i.e. pericentric regions or certain autosomal loci) are relatively low. Notably, as previously mentioned, SPT is a concentration independent technique and, thus, the presence of a bound population reflects solely the physical nature of binding and any changes in the amount of molecules bound are not correlated with the level of protein accumulation. Thus, it seems that the SmcHD1^{E147A} mutant displays loss of specificity of interaction and, while its binding to pericentric domains or the Xi chromosome is not affected per se, it is also binding persistently to many other regions of the genome. The observations made in this study are in agreement with a previously reported subcellular fractionation experiment where a construct expressing SmcHD1^{E147A} was more enriched in the chromatin-bound compared to the soluble nuclear fraction, while the WT counterpart showed equal distribution between the two fractions (Brideau et al., 2015). However, assaying the level of X-linked gene silencing in SmcHD1^{E147A} mESCs taken through ML differentiation for 7 days showed gene repression is affected in a SmcHD1-dependent manner (Figure 5.12), which suggests that the mutant protein, although bound to the Xi, does not perform its normal function. This is probably due to the greatly decreased residence time of SmcHD1^{E147A} over the Xi territory, as shown by the FRAP assay (Figure 5.11 H). Moreover, this decrease probably contributes also to lack of SmcHD1^{E147A} enrichment at target sites (Xi and pericentric domains).

To confirm the latter speculation, it is important to consider also the effects of the Y353C mutation. The bound population in cells having SmcHD1^{Y353C} is not significantly affected nucleus-wide, displaying only a very modest decrease (Figure 5.4 B-E, 5.11). On the other hand, the lifetime of the mutant on chromatin, in the nucleoplasm, is severely reduced (Figure 5.11 H). These two observations, together with the observed lack of accumulation of SmcHD1^{Y353C} at pericentric regions, suggest that SmcHD1 needs to reside at its target domains for a sufficiently long time to show enrichment. Over the Xi, at day 7, in SmcHD1^{Y353C} cells, the lifetime is massively affected, more than in the SmcHD1^{E147A} mutant, showing similar characteristics as observed in the nucleoplasm (Figure 5.11 H). This causes impaired SmcHD1 function in XCI gene repression (Figure 5.12) and, to a lesser extent, chromosome compaction (Figure 5.13). The reduced level of SmcHD1 bound to the Xi at day 7 (Figure 5.10) contributes as well to these two functional defects.

In SmcHD1^{G137E} cells, as in SmcHD1^{E147A} cells, there is an increase in the bound population of the protein across the whole nucleus (Figure 5.3 B-D, 5.11). However, unlike SmcHD1^{E147A}, SmcHD1^{G137E} retains its binding specificity. The quantity of bound molecules does go up also in a random nuclear region, which can include pericentric and other H3K9me3-marked regions, but, importantly, not to the same level as over the Xi territory, as shown first by SPT (Figure 5.8) and confirmed by FRAP (Figure 5.11). Moreover, SmcHD1^{G137E}'s residence time is not impaired, but it is actually greatly increased both in the nucleoplasm and across the Xi, even though the latter is difficult to estimate as previously explained (Section 5.4, Figure 5.11). Therefore, enrichment of SmcHD1 over the pericentric (Figure 5.3 A) and the Xi (Figure 5.7) domains, as well as its role in X-linked gene silencing (Figure 5.12) are not impeded. Furthermore, the accelerated accumulation of SmcHD1^{G137E} over the Xi territory (Figure 5.7), enabled by its higher binding tendency (Figure 5.8 A, F), and probably together with the increased lifetime, favour faster gene repression (based on

preliminary data assessing day 5 of X-linked gene silencing, Figure 5.12). Moreover, the mutation leads to a decrease in the volume of the Xi-associated H3K27me3 domain (Figure 5.13). This may be either due to SmcHD1^{G137E} ensuring higher insulation against PRC2 as previously observed with the SmcHD1 *MommeD43* mutant, which functions primarily as a gain-of-function mutation, (Tapia Del Fierro et al., 2023) or due to an increase in the overall compaction of the Xi. To distinguish the latter from the former, it would be useful to carry out DNA FISH with probes for the X chromosome, using the eGFP-CIZ1 signal to identify the inactive over the active X chromosomes. One note to make is that it is slightly surprising that the bound population of SmcHD1^{G137E} within a random region at day 7 is not larger than observed and closer to the one present at day 0 in a random region (Figure 5.8 F), but the reason behind this is currently unknown.

Overall, this study shows that SmcHD1 mutants which affect the GHKL ATPase activity of the protein in *in vitro* assays, also affect the dynamic behaviour of SmcHD1 in an *in vivo* cellular system, ultimately impacting its localization and functions regarding Xi-associated gene repression and Xi architecture. This suggests that the ATPase activity of SmcHD1 is required for fulfilment of its role in XCI. A model of how each mutation's effect on the ATPase activity may translate into an effect on the protein's interaction with chromatin and function will be discussed in Chapter 6.

Chapter 6: General Discussion

6.1 Investigating the molecular mechanism of SmcHD1 using live imaging

This study reports a detailed analysis of how various mutations affect the nuclear wide and Xi-specific behaviour and functionality of SmcHD1, relying mostly on live-cell imaging-based techniques. Due to its higher sensitivity compared to the previously employed method of fixed cell immunostaining, the use of live imaging has proven to be crucial in determining aspects such as the enrichment of SmcHD1 at pericentric regions on a population-wide level rather than in a fraction of cells and the gradual, rather than a late all or nothing, accumulation of SmcHD1 over the Xi chromosome during XCI progression.

Notably, relying only on visual inspection of the effect that a certain mutant has on SmcHD1's nuclear distribution can be misleading in terms on how it impacts the protein's interaction with chromatin. For example, knocking out *Lrif1* or introducing the E147A or the Y353C mutation into SmcHD1 lead to very similar phenotypes, with impaired enrichment over pericentric regions and the Xi, but the reasons behind these effects are actually different in each case, so looking at the proteins dynamic behaviour is required to uncover them. The cell line generated in this project, combining endogenously Halo-tagged SmcDH1 and transgenic expression of eGFP-tagged CIZ1 from the *ROSA26* locus, together with the Xi-specific SPT analysis setup allowed for an in-depth investigation into the dynamic behaviour of WT/mutant SmcHD1 not only across the whole nucleus, but also specifically over the Xi chromosome. This combination is key for being able to pair the

kinetic information corresponding to the Xi genomic location with data collected from other assays looking at Xi features, thus, gaining a comprehensive understanding of the characteristic behaviour of SmcHD1. Moreover, it is important to note that a setup such as the one employed in this project could, in the future, be used for the study of other factors involved in X chromosome silencing, as well as factors acting at a different nuclear region which can be highlighted in live imaging by a taggable mark. There is a limitation of the Xi-specific pipeline in that the relatively small number of analysable molecules, as discussed in section 4.3, precludes the determination of a residence time in the selected area in the case of SmcHD1. Such information can however be extracted using FRAP assays, as applied here. Alternatively, in some scenarios it may not be an issue if the nuclear ROI is larger than the Xi and, hence, would contain more molecules.

6.2 Two distinct localization pathways contribute to SmcHD1 enrichment over the Xi

Visualization of SmcHD1 and LRIF1 enrichment at the pericentric regions, coupled with the former's dramatically impaired accumulation in *Lrifi*^{-/-} mESCs (Figures 4.2, 4.8) cements the role of the LRIF1-HP1-H3K9me3 pathway in mediating SmcHD1's interaction with these heterochromatic domains. Moreover, as shown by SPT across the whole nucleus, this localization pathway accounts for almost all chromatin-bound SmcHD1, with absence of LRIF1 causing a marked redistribution of SmcHD1 from the bound to the fast-moving state, in both undifferentiated and differentiated mESCs (Figure 4.3). This result highlights the key role of LRIF1 in facilitating binding of SmcHD1 to chromatin. Apart pericentric regions, this pathway probably facilitates also association of SmcHD1 with other H3K9me3 sites, such as transposable elements and some repressed genes, along chromosome arms.

Moreover, it is worth mentioning that the chromodomain of HP1 can also bind H3K9me2 (Jacobs & Khorasanizadeh, 2002; Nielsen et al., 2002), so perhaps the LRIF1-mediated pathway ensures interaction of SmcHD1 with chromatin environments containing this mark as well. Like H3K9me3, H3K9me2 is associated with heterochromatic regions and plays a role in transcriptional silencing (reviewed in (Padeken et al., 2022)).

It was, however, surprising to find that LRIF1 is required for accumulation of SmcHD1 over the Xi chromosome in mESCs taken through ML differentiation, as silencing is being established, (Figure 4.4-6), given prior work in our lab showing that loss of LRIF1 does not affect SmcHD1 localisation to Xi in MEFs (Brideau et al., 2015). Moreover, in unpublished work we have observed LRIF1 to be dispensable for Xi enrichment in mESCs undergoing NPC differentiation (Dr. Mafalda Almeida, personal communication). In the NPC model, differentiating *Lrifi*^{-/-} mESCs did not present lack of enrichment of SmcHD1 over the Xi domain by immunostaining or defective X-linked gene silencing. Unfortunately, it was not possible to perform SPT experiments for SmcHD1 in NPCs due to a high autofluorescence level, potentially due to a metabolite produced in these cells or an additive in the NPC media. Nevertheless, the results obtained for NPC differentiation using conventional imaging, demonstrate that cells are able to compensate for the absence of LRIF1.

There are two aspects which should be considered regarding the apparently different involvement of LRIF1 in the recruitment of SmcHD1 to the Xi in ML compared to NPC differentiation. Firstly, in LRIF1 KO cells, the bound population of SmcHD1 within the Xi territory is already greatly diminished relative to WT cells at day 0 (Figure 4.5 A). This suggests that there is a pool of bound SmcHD1 on the future Xi, presumably dependent on the pre-existing H3K9me3. Secondly, enrichment of SmcHD1 on the Xi depends on differentiation, though it is not known exactly how, and on the presence of the H2AK119ub1

chromatin mark over the domain as previously shown and then confirmed in this study (Figures 3.2, 4.9, 4.10) (Bowness et al., 2022; Jansz, Nesterova, et al., 2018). However, both an immunoprecipitation assay in cells and a H2AK119ub1-marked nucleosome pulldown experiment have shown SmcHD1 does not directly bind to H2AK119ub1 (Dr. Mafalda Almeida, personal communication) (Jansz, Nesterova, et al., 2018). One could envision that the LRIF1-mediated interaction with the pre-XCI H3K9me3 is required for an initial localization and binding of SmcHD1 across the Xi, which then allows SmcHD1 to find the differentiation-dependent factor(s) (factor X) that enables the indirect interaction with H2AK119ub1 and gradual accumulation of SmcHD1 (Figure 6.1). It may be that factor X is more lowly expressed, at least in these early stages, in ML compared to NPC differentiation. Thus, when LRIF1 is absent in cells undergoing ML differentiation, SmcHD1 is unable to switch to the H2AK119ub1-dependent Xi enrichment pathway, while in NPCs, the more highly expressed factor X can compensate for the lack of the initial LRIF1-mediated binding of SmcHD1 to the Xi. In favour of a difference in expression of factor X due to the differentiation protocol, it is worth pointing out that, in the ML differentiation, the mesoderm and endoderm lineages are the ones represented (Heath & Smith, 1988; Heo et al., 2005), while NPCs are part of the ectoderm lineage.

The handover from the LRIF1-dependent to the H2AK119ub1-dependent pathway may be necessary, because H3K9me3 does not have a high abundance and coverage across the chromosome as H2AK119ub1 does (Bowness et al., 2022; de Napoles et al., 2004; Ichihara et al., 2022; Keniry et al., 2016; Rens et al., 2010) and, thus, SmcHD1 would not be able to reach a high enough number of potential binding sites in order to fulfil its function. If this model is true, one would expect that, upon looking at SmcHD1's dynamics in the Xist^{APID} cell line presented above, the bound population of the protein would not be affected over the Xi territory at day 0, but only at day 7. To determine exactly when the switch from the

LRIF1- to the H2AK119ub1-dependent pathway takes place, it may be interesting to employ a degron-based system to acutely remove LRIF1 at various points throughout the differentiation and assess its impact on SmcHD1 enrichment over the Xi domain.

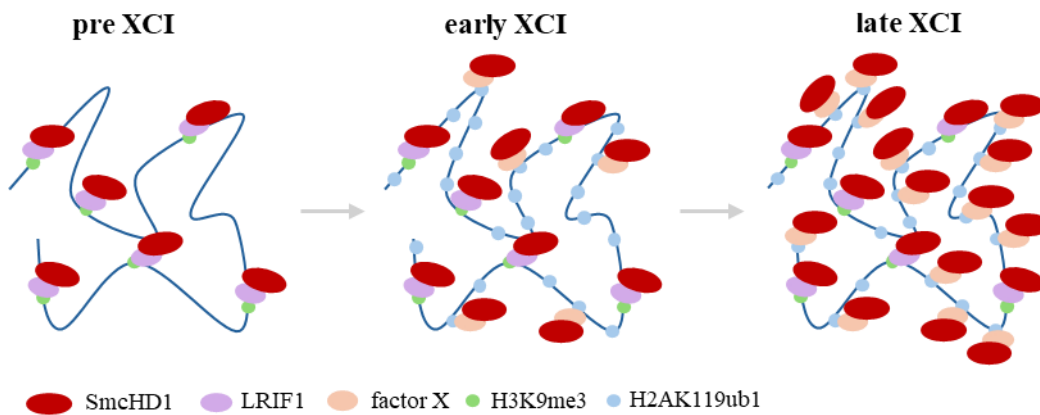


Figure 6.1 – Model of H3K9me3- and H2AK119ub1-dependent recruitment of SmcHD1 to Xi:

Before the onset of XCI coupled with cellular differentiation, SmcHD1 binds to the X chromosomes through an LRIF1-mediated association with H3K9me3. Once XCI starts to progress, on the chosen future Xi, H2AK119ub1 is deposited by PRC1. A differentiation-dependent factor (factor X) enables the SmcHD1 molecules already bound to the chromosome to associate with this abundant histone mark and, hence, gradually accumulate over the Xi in order to perform its functions in gene silencing, chromosomal architecture and CGI methylation.

6.3 The ATPase activity of SmcHD1 is required for its proper interaction with chromatin and functionality in XCI

The Xi-focused SPT analysis revealed SmcHD1 displays altered dynamics over the Xi territory compared to the rest of the nucleus, presenting a larger bound population which is

in accordance with the association being required to allow SmcHD1 to perform its function in XCI (Figure 3.4). Notably, as XCI progresses and SmcHD1 gradually accumulates over the domain and starts to mediate processes such as gene silencing, the bound population also increases (Figure 3.4). Moreover, the residence time of the protein within the Xi territory, as assayed by FRAP at day 7 of ML differentiation, is very high, indicating a very stable interaction with the underlying chromatin (Figure 5.11). The residence time is much larger than that of previously analysed XCI factors, SPEN and the PRC1 subunits Pcgf3/5, which, as discussed in Section 1.2.2.2, also mediate X-linked gene silencing (Almeida et al., 2017; Markaki et al., 2021). This suggests the mechanism of action employed by SmcHD1 to ensure gene repression might be quite different from that of these other proteins.

To gain more insight into this mechanism and into how SmcHD1 interacts with the underlying chromatin, a series of mutants was studied, specifically mutations affecting the ATPase activity of SmcHD1 previously reported to have a role in ensuring enrichment of the protein over the Xi chromosome (Brideau et al., 2015) as well in SmcHD1-linked disease-phenotypes (Jansz et al., 2017). Thus, two of the chosen mutations, G137E and Y353C, are found in BAMS and FSHD2 patients, respectively. Studying the consequences of such mutations on the functionality of SmcHD1 is also important from the perspective of needing to understand better the pathogenesis of the two disorders and, subsequently, find a targeted therapeutic approach, which currently does not exist in either case. As noted above, FSHD2 and BAMS occur as a consequence of heterozygous SmcHD1 mutations, indicating that the abrogation of D4Z4 repression is caused by WT-mutant heterodimers and/or mutant-mutant homodimers. The third analysed mutation which affects the catalytic residue 147, E147A, is not found in patients, but is of interest as it retains ATP binding but abolishes ATP hydrolysis, similar to the highly studied cohesin EQ mutant (Smc1 E1158Q/Smc3 E1155Q). The EQ mutation in either Smc1 or Smc3 does not impact formation of the tripartite complex Smc1-

Smc3-Scc1, and similarly E147A does not impair homodimerization of SmcHD1 (Arumugam et al., 2003; Brideau et al., 2015). While the exact mechanism for chromatin loading/unloading and function of the canonical Smc complex cohesin is most likely quite distinct from that of SmcHD1, given their structural differences and the many regulatory partners that cohesin relies on, it is interesting to note that the EQ mutation prevents stable cohesin loading onto chromosomes, highlighting the importance of ATP hydrolysis for this association (Arumugam et al., 2003; Hu et al., 2011). As seen here for SmcHD1^{E147A}, FRAP assays have also shown that EQ mutants have a faster turn over at target sites compared to their WT counterparts (Hu et al., 2011; Ladurner et al., 2014).

Integrating all of the information gained from the analysis of these three selected mutants, supports that, as previously hypothesized, SmcHD1 could function similarly to other GHKL ATPases via a molecular clamp mechanism mediated by the ATP binding-hydrolysis-release cycle (Dutta & Inouye, 2000; Gurzau et al., 2021; Pedersen et al., 2019) (Figure 6.2 A). As suggested by structural studies, ATP binding probably leads to dimerization of the two N-terminal heads of the SmcHD1 homodimer (Gurzau et al., 2021; Pedersen et al., 2019), with this conformational change favouring DNA binding as well as ATP hydrolysis. It is unclear where exactly DNA binding would take place and the structure of dimeric full length SmcHD1 will probably shed light on this aspect. One candidate could be the N-terminus, where a cavity appears when dimerization of the ATPase domains takes place, but DNA does not seem to be able to fit due to steric and electrostatic clashes (Pedersen et al., 2019). The remaining candidate is, thus, the hinge domain (Brideau et al., 2015; Chen et al., 2020; Chen et al., 2015). Once bound to chromatin, ATP hydrolysis would take place and, since *in vitro* studies have shown that SmcHD1 only has a weak ATPase activity, similar to that of other GHKL ATPases (Gurzau et al., 2021), the reaction is likely not able to power a motor function such as the loop extrusion process carried out by cohesin.

However, it may drive conformational changes within the protein itself, as seen, for example, for the ATPase domain of Type IIA DNA topoisomerase (Wei et al., 2005). One can speculate that such a conformational change could then allow interaction with another SmcHD1 molecule bound to a different location, allowing, thus, for formation of long-range chromatin interactions (Figure 6.2 B). While there is no evidence that such an interaction between SmcHD1 dimers takes place, proteomic studies have also not revealed any other possible candidates that SmcHD1 may interact with to fulfil its functions (Brideau et al., 2015). This type of SmcHD1-SmcHD1 interaction could underpin the stably localised molecules that account for the enrichment of the protein at genomic target regions such as the Xi. The last step of the cycle would then be ADP release and release of SmcHD1 from chromatin.

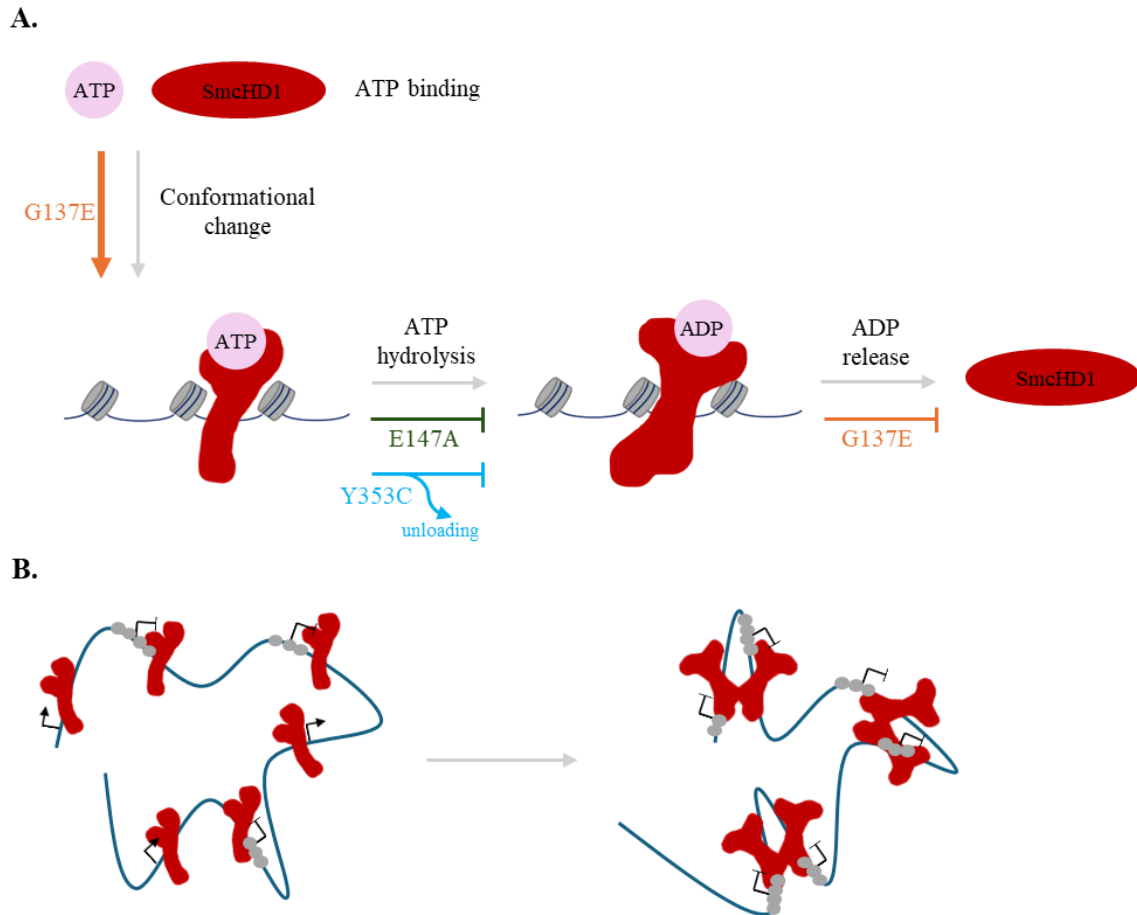


Figure 6.2 – A model for SmcHD1 interaction with chromatin and the Xi:

The model invokes that ATP binding to SmcHD1 leads to a conformational change that includes dimerization of the N-terminal GHKL ATPase domains, favours ATP hydrolysis and enables initial loading onto chromatin. Subsequently, ATP hydrolysis takes places, leading to a new protein conformation which favours stable and specific association with the chromatin and, potentially, allows distant DNA-bound SmcHD1 molecules to associate with one another (as shown in B.), forming long range chromatin interactions. Once ADP is released, SmcHD1 unloads from chromatin. The G137E mutation favours ATP binding and, hence, SmcHD1 loading onto chromatin and may also disfavour ADP release and, thus, SmcHD1 falling off chromatin. The E147A mutation prevents ATP hydrolysis and blocks the protein in the less stably chromatin-associated ATP-bound state. The Y353C mutation blocks the catalytic reaction and causes SmcHD1 to unload from chromatin. **B.** In the context of the Xi, the formation of SmcHD1-mediated long-range chromatin interactions may enable spreading of heterochromatic epigenetic marks and silencing factors and, consequently, gene silencing.

The G137E mutation probably causes a default conformation that favours ATP binding, given that N-terminal SmcHD1 constructs harbouring this mutation display increased ADP production and dimerization capacity *in vitro* (Gurzau et al., 2018; Pedersen et al., 2019). This translates into the observed increase in the bound population of SmcHD1 observed by SPT and FRAP (Figure 5.3, 5.8, 5.11). Moreover, the increase in the residence time on chromatin (Figure 5.11) suggests the G137E-determined conformation may also disfavour the release of ADP and, thus, the unloading of SmcHD1 from chromatin. SmcHD1^{G137E} functions well in the context of XCI, ensuring gene repression, which, if anything, is slightly accelerated (Figure 5.12). This is intriguing given the association of this mutation with D4Z4 hypomethylation in BAMS patients (Shaw et al., 2017). In this respect, it would, hence, be interesting to assess the Xi DNA methylation level in the cellular system presented here, as well as gene silencing at a later timepoint. It may be that the apparently higher efficiency of SmcHD1^{G137E} causes a de-synchronization with other XCI-pathways/players including those that link it to DNA methylation, and this may impair maintenance of the repressed state of X-linked genes, despite their initial silencing.

Regarding the E147A mutant, its retained ability to bind ATP (Pedersen et al., 2019) allows SmcHD1 to load onto chromatin, but, since it can't perform the hydrolysis reaction, it remains blocked in this state causing the observed increase in the bound population (Figure 5.2, 5.6, 5.11). Importantly, unlike SmcHD1^{G137E} which retains its binding specificity to target locations such as pericentric domains and the Xi (Figure 5.3, 5.7, 5.8), SmcHD1^{E147A} is not able to accumulate or reside for long enough at these regions and seems to be non-specifically bound across the genome (Figure 5.2, 5.5, 5.6, 5.11), stuck in a semi-engaged state. This indicates that the structural state reached via the ATP-induced conformational change is required for SmcHD1's recognition of and stable association with the target sites.

SmcHD1^{E147A}, thus, cannot fulfil its function in X-linked chromosome silencing (Figure 5.12).

The Y353C mutation abolishes the ATPase activity of SmcHD1 *in vitro* (Gurzau et al., 2018), and it may do so either by impairing ATP binding or the catalytic reaction itself. The very modest decrease in the bound population of SmcHD1^{Y353C} over the whole nucleus (Figure 5.4) shows engagement with chromatin is not severely impaired and, hence, argues against the former option. The dramatic decrease in the residence time of SmcHD1^{Y353C}, even lower than for SmcHD1^{E147A} (Figure 5.11), suggests that, despite the initial loading, SmcHD1^{Y353C} cannot initiate ATP hydrolysis and very quickly falls off the chromatin. This prevents the protein from enriching at pericentric regions or the Xi (Figure 5.4, 5.9) and performing its functions in XCI (Figure 5.12, 5.13). Similarly, at the D4Z4 array, SMCHD1 harbouring the Y353C mutation is probably unable to stably engage with the underlying chromatin and, thus, ensure repression of the repeat, leading to onset of FSHD2.

Overall, the mutants studied here shed light on the key role that the ATPase activity of SmcHD1 has on its molecular mechanism of interaction with the underlying chromatin and, consequently, on its functionality. Analysing other mutants as well as pairing this information with *in vitro* structural and functional analysis of full-length SmcHD1 will allow for a comprehensive understanding of SmcHD1's physiological and pathological roles.

Appendix

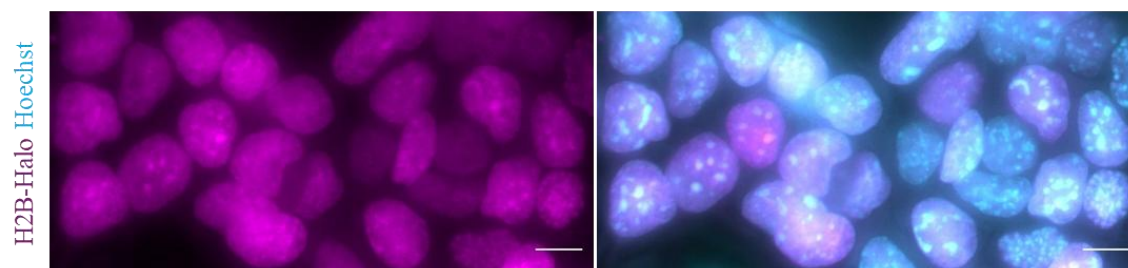


Figure A1 – Histone H2B-Halo mESC line:

Representative live images of Halo-tag labelling and Hoechst staining in histone H2B-Halo mESC line. Images are maximum-intensity projections, scale bar corresponds to 10 μ m.

References

- Aladjem, M. I., & Fu, H. (2014). A new light on DNA replication from the inactive X chromosome. *Bioessays*, 36(6), 591-597. <https://doi.org/10.1002/bies.201400021>
- Almeida, M., Bowness, J. S., & Brockdorff, N. (2020). The many faces of Polycomb regulation by RNA. *Curr Opin Genet Dev*, 61, 53-61. <https://doi.org/10.1016/j.gde.2020.02.023>
- Almeida, M., Pintacuda, G., Masui, O., Koseki, Y., Gdula, M., Cerase, A., Brown, D., Mould, A., Innocent, C., Nakayama, M., Schermelleh, L., Nesterova, T. B., Koseki, H., & Brockdorff, N. (2017). PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science*, 356(6342), 1081-1084. <https://doi.org/10.1126/science.aal2512>
- Ariyoshi, M., & Schwabe, J. W. (2003). A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. *Genes Dev*, 17(15), 1909-1920. <https://doi.org/10.1101/gad.266203>
- Arumugam, P., Gruber, S., Tanaka, K., Haering, C. H., Mechtler, K., & Nasmyth, K. (2003). ATP hydrolysis is required for cohesin's association with chromosomes. *Curr Biol*, 13(22), 1941-1953. <https://doi.org/10.1016/j.cub.2003.10.036>
- Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res*, 21(3), 381-395. <https://doi.org/10.1038/cr.2011.22>
- Barr, M. L., & Bertram, E. G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, 163(4148), 676. <https://doi.org/10.1038/163676a0>
- Benetti, N., Gouil, Q., Tapia Del Fierro, A., Beck, T., Breslin, K., Keniry, A., McGlinn, E., & Blewitt, M. E. (2022). Maternal SMCHD1 regulates Hox gene expression and patterning in the mouse embryo. *Nat Commun*, 13(1), 4295. <https://doi.org/10.1038/s41467-022-32057-x>
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., & Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol Cell*, 2(4), 437-445. [https://doi.org/10.1016/s1097-2765\(00\)80143-4](https://doi.org/10.1016/s1097-2765(00)80143-4)
- Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., McGouran, J. F., Hanssen, L. L. P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H. K., Sheahan, T. W., Brockdorff, N., Kessler, B. M., Koseki, H., & Klose, R. J. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*, 157(6), 1445-1459. <https://doi.org/10.1016/j.cell.2014.05.004>
- Blackledge, N. P., & Klose, R. J. (2021). The molecular principles of gene regulation by Polycomb repressive complexes. *Nat Rev Mol Cell Biol*, 22(12), 815-833. <https://doi.org/10.1038/s41580-021-00398-y>
- Blewitt, M. E., Gendrel, A. V., Pang, Z., Sparrow, D. B., Whitelaw, N., Craig, J. M., Apedaile, A., Hilton, D. J., Dunwoodie, S. L., Brockdorff, N., Kay, G. F., & Whitelaw, E. (2008). SmcHD1, containing a structural-maintenance-of-

- chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet*, 40(5), 663-669. <https://doi.org/10.1038/ng.142>
- Blewitt, M. E., Vickaryous, N. K., Hemley, S. J., Ashe, A., Bruxner, T. J., Preis, J. I., Arkell, R., & Whitelaw, E. (2005). An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proc Natl Acad Sci U S A*, 102(21), 7629-7634. <https://doi.org/10.1073/pnas.0409375102>
- Bowness, J. S., Nesterova, T. B., Wei, G., Rodermund, L., Almeida, M., Coker, H., Carter, E. J., Kadaster, A., & Brockdorff, N. (2022). Xist-mediated silencing requires additive functions of SPEN and Polycomb together with differentiation-dependent recruitment of SmcHD1. *Cell Rep*, 39(7), 110830. <https://doi.org/10.1016/j.celrep.2022.110830>
- Brideau, N. J., Coker, H., Gendrel, A. V., Siebert, C. A., Bezstarosti, K., Demmers, J., Poot, R. A., Nesterova, T. B., & Brockdorff, N. (2015). Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. *Mol Cell Biol*, 35(23), 4053-4068. <https://doi.org/10.1128/MCB.00432-15>
- Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., Norris, D. P., Penny, G. D., Patel, D., & Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature*, 351(6324), 329-331. <https://doi.org/10.1038/351329a0>
- Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S., & Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, 71(3), 515-526. [https://doi.org/10.1016/0092-8674\(92\)90519-i](https://doi.org/10.1016/0092-8674(92)90519-i)
- Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J., & Willard, H. F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, 71(3), 527-542. [https://doi.org/10.1016/0092-8674\(92\)90520-m](https://doi.org/10.1016/0092-8674(92)90520-m)
- Brown, S. D. (1991). XIST and the mapping of the X chromosome inactivation centre. *Bioessays*, 13(11), 607-612. <https://doi.org/10.1002/bies.950131112>
- Cavalli, G., & Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. *Nature*, 571(7766), 489-499. <https://doi.org/10.1038/s41586-019-1411-0>
- Cerase, A., Armaos, A., Neumayer, C., Avner, P., Guttman, M., & Tartaglia, G. G. (2019). Phase separation drives X-chromosome inactivation: a hypothesis. *Nat Struct Mol Biol*, 26(5), 331-334. <https://doi.org/10.1038/s41594-019-0223-0>
- Cerase, A., Smeets, D., Tang, Y. A., Gdula, M., Kraus, F., Spivakov, M., Moindrot, B., Leleu, M., Tattermusch, A., Demmerle, J., Nesterova, T. B., Green, C., Otte, A. P., Schermelleh, L., & Brockdorff, N. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. *Proc Natl Acad Sci U S A*, 111(6), 2235-2240. <https://doi.org/10.1073/pnas.1312951111>
- Chadwick, B. P. (2008). DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. *Genome Res*, 18(8), 1259-1269. <https://doi.org/10.1101/gr.075713.107>
- Chaumeil, J., Le Baccon, P., Wutz, A., & Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are

- recruited when silenced. *Genes Dev*, 20(16), 2223-2237. <https://doi.org/10.1101/gad.380906>
- Chen, C. K., Blanco, M., Jackson, C., Aznauryan, E., Ollikainen, N., Surka, C., Chow, A., Cerase, A., McDonel, P., & Guttman, M. (2016). Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science*, 354(6311), 468-472. <https://doi.org/10.1126/science.aae0047>
- Chen, K., Birkinshaw, R. W., Gurzau, A. D., Wanigasuriya, I., Wang, R., Iminoff, M., Sandow, J. J., Young, S. N., Hennessy, P. J., Willson, T. A., Heckmann, D. A., Webb, A. I., Blewitt, M. E., Czabotar, P. E., & Murphy, J. M. (2020). Crystal structure of the hinge domain of Smchd1 reveals its dimerization mode and nucleic acid-binding residues. *Sci Signal*, 13(636). <https://doi.org/10.1126/scisignal.aaz5599>
- Chen, K., Czabotar, P. E., Blewitt, M. E., & Murphy, J. M. (2016). The hinge domain of the epigenetic repressor Smchd1 adopts an unconventional homodimeric configuration. *Biochem J*, 473(6), 733-742. <https://doi.org/10.1042/BJ20151049>
- Chen, K., Dobson, R. C., Lucet, I. S., Young, S. N., Pearce, F. G., Blewitt, M. E., & Murphy, J. M. (2016). The epigenetic regulator Smchd1 contains a functional GHKL-type ATPase domain. *Biochem J*, 473(12), 1733-1744. <https://doi.org/10.1042/BCJ20160189>
- Chen, K., Hu, J., Moore, D. L., Liu, R., Kessans, S. A., Breslin, K., Lucet, I. S., Keniry, A., Leong, H. S., Parish, C. L., Hilton, D. J., Lemmers, R. J., van der Maarel, S. M., Czabotar, P. E., Dobson, R. C., Ritchie, M. E., Kay, G. F., Murphy, J. M., & Blewitt, M. E. (2015). Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *Proc Natl Acad Sci U S A*, 112(27), E3535-3544. <https://doi.org/10.1073/pnas.1504232112>
- Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M., & Ge, K. (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem*, 282(28), 20395-20406. <https://doi.org/10.1074/jbc.M701574200>
- Chu, C., Zhang, Q. C., da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., & Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. *Cell*, 161(2), 404-416. <https://doi.org/10.1016/j.cell.2015.03.025>
- Chureau, C., Chantalat, S., Romito, A., Galvani, A., Duret, L., Avner, P., & Rougeulle, C. (2011). Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum Mol Genet*, 20(4), 705-718. <https://doi.org/10.1093/hmg/ddq516>
- Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., & Nasmyth, K. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell*, 5(2), 243-254. [https://doi.org/10.1016/s1097-2765\(00\)80420-7](https://doi.org/10.1016/s1097-2765(00)80420-7)
- Clapp, J., Mitchell, L. M., Bolland, D. J., Fantes, J., Corcoran, A. E., Scotting, P. J., Armour, J. A., & Hewitt, J. E. (2007). Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet*, 81(2), 264-279. <https://doi.org/10.1086/519311>

- Clemson, C. M., McNeil, J. A., Willard, H. F., & Lawrence, J. B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol*, 132(3), 259-275. <https://doi.org/10.1083/jcb.132.3.259>
- Coker, H., Wei, G., Moindrot, B., Mohammed, S., Nesterova, T., & Brockdorff, N. (2020). The role of the Xist 5' m6A region and RBM15 in X chromosome inactivation. *Wellcome Open Res*, 5, 31. <https://doi.org/10.12688/wellcomeopenres.15711.1>
- Collombet, S., Rall, I., Dugast-Darzacq, C., Heckert, A., Halavatyi, A., Le Saux, A., Dailey, G., Darzacq, X., & Heard, E. (2023). RNA polymerase II depletion from the inactive X chromosome territory is not mediated by physical compartmentalization. *Nat Struct Mol Biol*, 30(8), 1216-1223. <https://doi.org/10.1038/s41594-023-01008-5>
- Collombet, S., Ranisavljevic, N., Nagano, T., Varnai, C., Shisode, T., Leung, W., Piolot, T., Galupa, R., Borensztein, M., Servant, N., Fraser, P., Ancelin, K., & Heard, E. (2020). Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature*, 580(7801), 142-146. <https://doi.org/10.1038/s41586-020-2125-z>
- Colognori, D., Sunwoo, H., Wang, D., Wang, C. Y., & Lee, J. T. (2020). Xist Repeats A and B Account for Two Distinct Phases of X Inactivation Establishment. *Dev Cell*, 54(1), 21-32 e25. <https://doi.org/10.1016/j.devcel.2020.05.021>
- Concordet, J. P., & Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*, 46(W1), W242-W245. <https://doi.org/10.1093/nar/gky354>
- Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N. P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., Heger, A., & Brockdorff, N. (2014). Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep*, 7(5), 1456-1470. <https://doi.org/10.1016/j.celrep.2014.04.012>
- Cooper, S., Grijzenhout, A., Underwood, E., Ancelin, K., Zhang, T., Nesterova, T. B., Anil-Kirmizitas, B., Bassett, A., Kooistra, S. M., Agger, K., Helin, K., Heard, E., & Brockdorff, N. (2016). Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat Commun*, 7, 13661. <https://doi.org/10.1038/ncomms13661>
- Corsi, F., Rusch, E., & Goloborodko, A. (2023). Loop extrusion rules: the next generation. *Curr Opin Genet Dev*, 81, 102061. <https://doi.org/10.1016/j.gde.2023.102061>
- Csankovszki, G., Nagy, A., & Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J Cell Biol*, 153(4), 773-784. <https://doi.org/10.1083/jcb.153.4.773>
- Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R., & Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat Genet*, 22(4), 323-324. <https://doi.org/10.1038/11887>
- Darrow, E. M., Huntley, M. H., Dudchenko, O., Stamenova, E. K., Durand, N. C., Sun, Z., Huang, S. C., Sanborn, A. L., Machol, I., Shamim, M., Seberg, A. P., Lander, E. S., Chadwick, B. P., & Aiden, E. L. (2016). Deletion of DXZ4 on the human

- inactive X chromosome alters higher-order genome architecture. *Proc Natl Acad Sci U S A*, 113(31), E4504-4512. <https://doi.org/10.1073/pnas.1609643113>
- Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J. M. (2019). DNA loop extrusion by human cohesin. *Science*, 366(6471), 1338-1345. <https://doi.org/10.1126/science.aaz3418>
- Davidson, I. F., & Peters, J. M. (2021). Genome folding through loop extrusion by SMC complexes. *Nat Rev Mol Cell Biol*, 22(7), 445-464. <https://doi.org/10.1038/s41580-021-00349-7>
- de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H., & Brockdorff, N. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell*, 7(5), 663-676. <https://doi.org/10.1016/j.devcel.2004.10.005>
- Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev*, 25(10), 1010-1022. <https://doi.org/10.1101/gad.2037511>
- Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J. B., Blau, C. A., Shendure, J., Duan, Z., Noble, W. S., & Disteche, C. M. (2015). Bipartite structure of the inactive mouse X chromosome. *Genome Biol*, 16(1), 152. <https://doi.org/10.1186/s13059-015-0728-8>
- Dion, C., Roche, S., Laberthonniere, C., Broucqsault, N., Mariot, V., Xue, S., Gurzau, A. D., Nowak, A., Gordon, C. T., Gaillard, M. C., El-Yazidi, C., Thomas, M., Schlupp-Robaglia, A., Missirian, C., Malan, V., Ratbi, L., Sefiani, A., Wollnik, B., Binetruy, B., . . . Magdinier, F. (2019). SMCHD1 is involved in de novo methylation of the DUX4-encoding D4Z4 macrosatellite. *Nucleic Acids Res*, 47(6), 2822-2839. <https://doi.org/10.1093/nar/gkz005>
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398), 376-380. <https://doi.org/10.1038/nature11082>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21. <https://doi.org/10.1093/bioinformatics/bts635>
- Donohoe, M. E., Silva, S. S., Pinter, S. F., Xu, N., & Lee, J. T. (2009). The pluripotency factor Oct4 interacts with Ctf and also controls X-chromosome pairing and counting. *Nature*, 460(7251), 128-132. <https://doi.org/10.1038/nature08098>
- Dossin, F., Pinheiro, I., Zylitz, J. J., Roensch, J., Collombet, S., Le Saux, A., Chelmsicki, T., Attia, M., Kapoor, V., Zhan, Y., Dingli, F., Loew, D., Mercher, T., Dekker, J., & Heard, E. (2020). SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature*, 578(7795), 455-460. <https://doi.org/10.1038/s41586-020-1974-9>
- Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L., & Roeder, R. G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell*, 121(6), 873-885. <https://doi.org/10.1016/j.cell.2005.04.031>

- Duthie, S. M., Nesterova, T. B., Formstone, E. J., Keohane, A. M., Turner, B. M., Zakian, S. M., & Brockdorff, N. (1999). Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis. *Hum Mol Genet*, 8(2), 195-204. <https://doi.org/10.1093/hmg/8.2.195>
- Dutta, R., & Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci*, 25(1), 24-28. [https://doi.org/10.1016/s0968-0004\(99\)01503-0](https://doi.org/10.1016/s0968-0004(99)01503-0)
- Eberl, H. C., Spruijt, C. G., Kelstrup, C. D., Vermeulen, M., & Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell*, 49(2), 368-378. <https://doi.org/10.1016/j.molcel.2012.10.026>
- Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E. S., Plath, K., & Guttman, M. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science*, 341(6147), 1237973. <https://doi.org/10.1126/science.1237973>
- Escamilla-Del-Arenal, M., da Rocha, S. T., Spruijt, C. G., Masui, O., Renaud, O., Smits, A. H., Margueron, R., Vermeulen, M., & Heard, E. (2013). Cdy1, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. *Mol Cell Biol*, 33(24), 5005-5020. <https://doi.org/10.1128/MCB.00866-13>
- Fang, J., Chen, T., Chadwick, B., Li, E., & Zhang, Y. (2004). Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J Biol Chem*, 279(51), 52812-52815. <https://doi.org/10.1074/jbc.C400493200>
- Farcas, A. M., Blackledge, N. P., Sudbery, I., Long, H. K., McGouran, J. F., Rose, N. R., Lee, S., Sims, D., Cerase, A., Sheahan, T. W., Koseki, H., Brockdorff, N., Ponting, C. P., Kessler, B. M., & Klose, R. J. (2012). KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife*, 1, e00205. <https://doi.org/10.7554/eLife.00205>
- Fraser, P., & Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature*, 447(7143), 413-417. <https://doi.org/10.1038/nature05916>
- Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep*, 15(9), 2038-2049. <https://doi.org/10.1016/j.celrep.2016.04.085>
- Gdula, M. R., Nesterova, T. B., Pintacuda, G., Godwin, J., Zhan, Y., Ozadam, H., McClellan, M., Moralli, D., Krueger, F., Green, C. M., Reik, W., Kriaucionis, S., Heard, E., Dekker, J., & Brockdorff, N. (2019). The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nat Commun*, 10(1), 30. <https://doi.org/10.1038/s41467-018-07907-2>
- Gendrel, A. V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y. A., Huntley, D., Montana, G., Taylor, S., Giannoulidou, E., Heard, E., Stancheva, I., & Brockdorff, N. (2012). SmcHD1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev Cell*, 23(2), 265-279. <https://doi.org/10.1016/j.devcel.2012.06.011>

- Gendrel, A. V., Tang, Y. A., Suzuki, M., Godwin, J., Nesterova, T. B., Grealley, J. M., Heard, E., & Brockdorff, N. (2013). Epigenetic functions of smchd1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol Cell Biol*, 33(16), 3150-3165. <https://doi.org/10.1128/MCB.00145-13>
- Geng, L. N., Yao, Z., Snider, L., Fong, A. P., Cech, J. N., Young, J. M., van der Maarel, S. M., Ruzzo, W. L., Gentleman, R. C., Tawil, R., & Tapscott, S. J. (2012). DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev Cell*, 22(1), 38-51. <https://doi.org/10.1016/j.devcel.2011.11.013>
- Gilbert, D. M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol*, 14(3), 377-383. [https://doi.org/10.1016/s0955-0674\(02\)00326-5](https://doi.org/10.1016/s0955-0674(02)00326-5)
- Giorgetti, L., Lajoie, B. R., Carter, A. C., Attia, M., Zhan, Y., Xu, J., Chen, C. J., Kaplan, N., Chang, H. Y., Heard, E., & Dekker, J. (2016). Structural organization of the inactive X chromosome in the mouse. *Nature*, 535(7613), 575-579. <https://doi.org/10.1038/nature18589>
- Gontan, C., Achame, E. M., Demmers, J., Barakat, T. S., Rentmeester, E., van, I. W., Grootegoed, J. A., & Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature*, 485(7398), 386-390. <https://doi.org/10.1038/nature11070>
- Gordon, C. T., Xue, S., Yigit, G., Filali, H., Chen, K., Rosin, N., Yoshiura, K. I., Oufadem, M., Beck, T. J., McGowan, R., Magee, A. C., Altmuller, J., Dion, C., Thiele, H., Gurzau, A. D., Nurnberg, P., Meschede, D., Muhlbauer, W., Okamoto, N., . . . Reversade, B. (2017). De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat Genet*, 49(2), 249-255. <https://doi.org/10.1038/ng.3765>
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268(5218), 1766-1769. <https://doi.org/10.1126/science.7792603>
- Graves, J. A. (1982). 5-azacytidine-induced re-expression of alleles on the inactive X chromosome in a hybrid mouse cell line. *Exp Cell Res*, 141(1), 99-105. [https://doi.org/10.1016/0014-4827\(82\)90072-6](https://doi.org/10.1016/0014-4827(82)90072-6)
- Greenberg, M. V. C., & Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*, 20(10), 590-607. <https://doi.org/10.1038/s41580-019-0159-6>
- Grimm, J. B., English, B. P., Choi, H., Muthusamy, A. K., Mehl, B. P., Dong, P., Brown, T. A., Lippincott-Schwartz, J., Liu, Z., Lionnet, T., & Lavis, L. D. (2016). Bright photoactivatable fluorophores for single-molecule imaging. *Nat Methods*, 13(12), 985-988. <https://doi.org/10.1038/nmeth.4034>
- Gurzau, A. D., Chen, K., Xue, S., Dai, W., Lucet, I. S., Ly, T. T. N., Reversade, B., Blewitt, M. E., & Murphy, J. M. (2018). FSHD2- and BAMS-associated mutations confer opposing effects on SMCHD1 function. *J Biol Chem*, 293(25), 9841-9853. <https://doi.org/10.1074/jbc.RA118.003104>
- Gurzau, A. D., Horne, C. R., Mok, Y. F., Iminoff, M., Willson, T. A., Young, S. N., Blewitt, M. E., & Murphy, J. M. (2021). SMCHD1's ubiquitin-like domain is required for N-terminal dimerization and chromatin localization. *Biochem J*, 478(13), 2555-2569. <https://doi.org/10.1042/BCJ20210278>

- Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., & Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev Cell*, 19(3), 469-476. <https://doi.org/10.1016/j.devcel.2010.08.006>
- Hayakawa, T., Haraguchi, T., Masumoto, H., & Hiraoka, Y. (2003). Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *J Cell Sci*, 116(Pt 16), 3327-3338. <https://doi.org/10.1242/jcs.00635>
- Heath, J. K., & Smith, A. G. (1988). Regulatory factors of embryonic stem cells. *J Cell Sci Suppl*, 10, 257-266. https://doi.org/10.1242/jcs.1988.supplement_10.18
- Heo, J., Lee, J. S., Chu, I. S., Takahama, Y., & Thorgeirsson, S. S. (2005). Spontaneous differentiation of mouse embryonic stem cells in vitro: characterization by global gene expression profiles. *Biochem Biophys Res Commun*, 332(4), 1061-1069. <https://doi.org/10.1016/j.bbrc.2005.04.173>
- Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol*, 7(5), 311-322. <https://doi.org/10.1038/nrm1909>
- Horakova, A. H., Calabrese, J. M., McLaughlin, C. R., Tremblay, D. C., Magnuson, T., & Chadwick, B. P. (2012). The mouse DXZ4 homolog retains Ctf binding and proximity to Pls3 despite substantial organizational differences compared to the primate macrosatellite. *Genome Biol*, 13(8), R70. <https://doi.org/10.1186/gb-2012-13-8-r70>
- Horsfield, J. A. (2023). Full circle: a brief history of cohesin and the regulation of gene expression. *FEBS J*, 290(7), 1670-1687. <https://doi.org/10.1111/febs.16362>
- Hsieh, T. S., Cattoglio, C., Slobodyanyuk, E., Hansen, A. S., Rando, O. J., Tjian, R., & Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol Cell*, 78(3), 539-553 e538. <https://doi.org/10.1016/j.molcel.2020.03.002>
- Hu, B., Itoh, T., Mishra, A., Katoh, Y., Chan, K. L., Upcher, W., Godlee, C., Roig, M. B., Shirahige, K., & Nasmyth, K. (2011). ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. *Curr Biol*, 21(1), 12-24. <https://doi.org/10.1016/j.cub.2010.12.004>
- Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., & Xu, R. M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science*, 312(5774), 748-751. <https://doi.org/10.1126/science.1125162>
- Huseyin, M. K., & Klose, R. J. (2021). Live-cell single particle tracking of PRC1 reveals a highly dynamic system with low target site occupancy. *Nat Commun*, 12(1), 887. <https://doi.org/10.1038/s41467-021-21130-6>
- Ichihara, S., Nagao, K., Sakaguchi, T., Obuse, C., & Sado, T. (2022). SmcHD1 underlies the formation of H3K9me3 blocks on the inactive X chromosome in mice. *Development*, 149(15). <https://doi.org/10.1242/dev.200864>
- Inoue, A., Jiang, L., Lu, F., & Zhang, Y. (2017). Genomic imprinting of Xist by maternal H3K27me3. *Genes Dev*, 31(19), 1927-1932. <https://doi.org/10.1101/gad.304113.117>
- Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C. M., Nakamura, T., Mazo, A., Eisenbach, L., & Canaani, E. (2007). Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol Cell Biol*, 27(5), 1889-1903. <https://doi.org/10.1128/MCB.01506-06>

- Jachowicz, J. W., Strehle, M., Banerjee, A. K., Blanco, M. R., Thai, J., & Guttman, M. (2022). Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat Struct Mol Biol*, 29(3), 239-249. <https://doi.org/10.1038/s41594-022-00739-1>
- Jacobs, S. A., & Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*, 295(5562), 2080-2083. <https://doi.org/10.1126/science.1069473>
- James, T. C., & Elgin, S. C. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol Cell Biol*, 6(11), 3862-3872. <https://doi.org/10.1128/mcb.6.11.3862-3872.1986>
- Janssen, A., Colmenares, S. U., & Karpen, G. H. (2018). Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol*, 34, 265-288. <https://doi.org/10.1146/annurev-cellbio-100617-062653>
- Jansz, N., Chen, K., Murphy, J. M., & Blewitt, M. E. (2017). The Epigenetic Regulator SMCHD1 in Development and Disease. *Trends Genet*, 33(4), 233-243. <https://doi.org/10.1016/j.tig.2017.01.007>
- Jansz, N., Keniry, A., Trussart, M., Bildsoe, H., Beck, T., Tonks, I. D., Mould, A. W., Hickey, P., Breslin, K., Iminoff, M., Ritchie, M. E., McGlinn, E., Kay, G. F., Murphy, J. M., & Blewitt, M. E. (2018). Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. *Nat Struct Mol Biol*, 25(9), 766-777. <https://doi.org/10.1038/s41594-018-0111-z>
- Jansz, N., Nesterova, T., Keniry, A., Iminoff, M., Hickey, P. F., Pintacuda, G., Masui, O., Kobelke, S., Geoghegan, N., Breslin, K. A., Willson, T. A., Rogers, K., Kay, G. F., Fox, A. H., Koseki, H., Brockdorff, N., Murphy, J. M., & Blewitt, M. E. (2018). Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. *Cell Rep*, 25(7), 1912-1923 e1919. <https://doi.org/10.1016/j.celrep.2018.10.044>
- Jonkers, I., Barakat, T. S., Achame, E. M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J. A., & Gribnau, J. (2009). RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell*, 139(5), 999-1011. <https://doi.org/10.1016/j.cell.2009.10.034>
- Keniry, A., Gearing, L. J., Jansz, N., Liu, J., Holik, A. Z., Hickey, P. F., Kinkel, S. A., Moore, D. L., Breslin, K., Chen, K., Liu, R., Phillips, C., Pakusch, M., Biben, C., Sheridan, J. M., Kile, B. T., Carmichael, C., Ritchie, M. E., Hilton, D. J., & Blewitt, M. E. (2016). Setdb1-mediated H3K9 methylation is enriched on the inactive X and plays a role in its epigenetic silencing. *Epigenetics Chromatin*, 9, 16. <https://doi.org/10.1186/s13072-016-0064-6>
- Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., & Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol*, 21(1), 86-89. <https://doi.org/10.1038/nbt765>
- Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J., & Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. *Science*, 366(6471), 1345-1349. <https://doi.org/10.1126/science.aaz4475>
- King, H. W., & Klose, R. J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *Elife*, 6. <https://doi.org/10.7554/eLife.22631>

- Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., & Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature*, 442(7100), 312-316. <https://doi.org/10.1038/nature04853>
- Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., & Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLoS Biol*, 2(7), E171. <https://doi.org/10.1371/journal.pbio.0020171>
- Kolpa, H. J., Fackelmayer, F. O., & Lawrence, J. B. (2016). SAF-A Requirement in Anchoring XIST RNA to Chromatin Varies in Transformed and Primary Cells. *Dev Cell*, 39(1), 9-10. <https://doi.org/10.1016/j.devcel.2016.09.021>
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science*, 184(4139), 868-871. <https://doi.org/10.1126/science.184.4139.868>
- Krueger, F., & Andrews, S. R. (2016). SNPsplite: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res*, 5, 1479. <https://doi.org/10.12688/f1000research.9037.2>
- Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C., Xie, X., Chi, A. S., Adli, M., Kasif, S., Ptaszek, L. M., Cowan, C. A., Lander, E. S., Koseki, H., & Bernstein, B. E. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet*, 4(10), e1000242. <https://doi.org/10.1371/journal.pgen.1000242>
- Ladurner, R., Bhaskara, V., Huis in 't Veld, P. J., Davidson, I. F., Kreidl, E., Petzold, G., & Peters, J. M. (2014). Cohesin's ATPase activity couples cohesin loading onto DNA with Smc3 acetylation. *Curr Biol*, 24(19), 2228-2237. <https://doi.org/10.1016/j.cub.2014.08.011>
- Lai, F., Orom, U. A., Cesaroni, M., Beringer, M., Taatjes, D. J., Blobel, G. A., & Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature*, 494(7438), 497-501. <https://doi.org/10.1038/nature11884>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9(4), 357-359. <https://doi.org/10.1038/nmeth.1923>
- Latty, S. L., Felce, J. H., Weimann, L., Lee, S. F., Davis, S. J., & Klennerman, D. (2015). Referenced Single-Molecule Measurements Differentiate between GPCR Oligomerization States. *Biophys J*, 109(9), 1798-1806. <https://doi.org/10.1016/j.bpj.2015.09.004>
- Lee, J. H., & Skalnik, D. G. (2005). CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *J Biol Chem*, 280(50), 41725-41731. <https://doi.org/10.1074/jbc.M508312200>
- Lee, J. H., Tate, C. M., You, J. S., & Skalnik, D. G. (2007). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J Biol Chem*, 282(18), 13419-13428. <https://doi.org/10.1074/jbc.M609809200>
- Lee, J. T., Davidow, L. S., & Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet*, 21(4), 400-404. <https://doi.org/10.1038/7734>

- Lee, J. T., & Jaenisch, R. (1997). Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature*, 386(6622), 275-279. <https://doi.org/10.1038/386275a0>
- Lee, J. T., & Lu, N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell*, 99(1), 47-57. [https://doi.org/10.1016/s0092-8674\(00\)80061-6](https://doi.org/10.1016/s0092-8674(00)80061-6)
- Lemmers, R., van der Stoep, N., Vliet, P. J. V., Moore, S. A., San Leon Granado, D., Johnson, K., Topf, A., Straub, V., Evangelista, T., Mozaffar, T., Kimonis, V., Shaw, N. D., Selvatici, R., Ferlini, A., Voermans, N., van Engelen, B., Sacconi, S., Tawil, R., Lamers, M., & van der Maarel, S. M. (2019). SMCHD1 mutation spectrum for facioscapulohumeral muscular dystrophy type 2 (FSHD2) and Bosma arhinia microphthalmia syndrome (BAMS) reveals disease-specific localisation of variants in the ATPase domain. *J Med Genet*, 56(10), 693-700. <https://doi.org/10.1136/jmedgenet-2019-106168>
- Lemmers, R. J., Tawil, R., Petek, L. M., Balog, J., Block, G. J., Santen, G. W., Amell, A. M., van der Vliet, P. J., Almomani, R., Straasheijm, K. R., Krom, Y. D., Klooster, R., Sun, Y., den Dunnen, J. T., Helmer, Q., Donlin-Smith, C. M., Padberg, G. W., van Engelen, B. G., de Greef, J. C., . . . van der Maarel, S. M. (2012). Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet*, 44(12), 1370-1374. <https://doi.org/10.1038/ng.2454>
- Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature*, 276(5688), 565-570. <https://doi.org/10.1038/276565a0>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, Y., Haarhuis, J. H. I., Sedenio Cacciatore, A., Oldenkamp, R., van Ruiten, M. S., Willems, L., Teunissen, H., Muir, K. W., de Wit, E., Rowland, B. D., & Panne, D. (2020). The structural basis for cohesin-CTCF-anchored loops. *Nature*, 578(7795), 472-476. <https://doi.org/10.1038/s41586-019-1910-z>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-930. <https://doi.org/10.1093/bioinformatics/btt656>
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950), 289-293. <https://doi.org/10.1126/science.1181369>
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., & Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*, 12(8), 767-772. <https://doi.org/10.1038/nmeth.3453>
- Loda, A., Collombet, S., & Heard, E. (2022). Gene regulation in time and space during X-chromosome inactivation. *Nat Rev Mol Cell Biol*, 23(4), 231-249. <https://doi.org/10.1038/s41580-021-00438-7>

- Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., Davidovich, C., Gooding, A. R., Goodrich, K. J., Mattick, J. S., Mesirov, J. P., Cech, T. R., & Chang, H. Y. (2016). RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell*, 165(5), 1267-1279. <https://doi.org/10.1016/j.cell.2016.04.028>
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature*, 389(6648), 251-260. <https://doi.org/10.1038/38444>
- Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*, 190, 372-373. <https://doi.org/10.1038/190372a0>
- Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J. I., Wolf, M., & Kurumizaka, H. (2018). Structural Basis of Heterochromatin Formation by Human HP1. *Mol Cell*, 69(3), 385-397 e388. <https://doi.org/10.1016/j.molcel.2017.12.011>
- Mak, W., Nesterova, T. B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A. P., & Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. *Science*, 303(5658), 666-669. <https://doi.org/10.1126/science.1092674>
- Markaki, Y., Chong, J. G., Wang, Y., Jacobson, E. C., Luong, C., Tan, S. Y. X., Jachowicz, J. W., Strehle, M., Maestrini, D., Banerjee, A. K., Mistry, B. A., Dror, I., Dossin, F., Schöneberg, J., Heard, E., Guttman, M., Chou, T., & Plath, K. (2021). Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell*, 184(25), 6212. <https://doi.org/10.1016/j.cell.2021.11.028>
- McHugh, C. A., Chen, C. K., Chow, A., Surka, C. F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., Sweredoski, M. J., Shishkin, A. A., Su, J., Lander, E. S., Hess, S., Plath, K., & Guttman, M. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), 232-236. <https://doi.org/10.1038/nature14443>
- Mei, H., Kozuka, C., Hayashi, R., Kumon, M., Koseki, H., & Inoue, A. (2021). H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. *Nat Genet*, 53(4), 539-550. <https://doi.org/10.1038/s41588-021-00820-3>
- Mermoud, J. E., Costanzi, C., Pehrson, J. R., & Brockdorff, N. (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J Cell Biol*, 147(7), 1399-1408. <https://doi.org/10.1083/jcb.147.7.1399>
- Milne, T. A., Kim, J., Wang, G. G., Stadler, S. C., Basrur, V., Whitcomb, S. J., Wang, Z., Ruthenburg, A. J., Elenitoba-Johnson, K. S., Roeder, R. G., & Allis, C. D. (2010). Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol Cell*, 38(6), 853-863. <https://doi.org/10.1016/j.molcel.2010.05.011>
- Minajigi, A., Froberg, J., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., & Lee, J. T. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science*, 349(6245). <https://doi.org/10.1126/science.aab2276>

- Minc, E., Allory, Y., Worman, H. J., Courvalin, J. C., & Buendia, B. (1999). Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma*, 108(4), 220-234. <https://doi.org/10.1007/s004120050372>
- Minc, E., Courvalin, J. C., & Buendia, B. (2000). HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet Cell Genet*, 90(3-4), 279-284. <https://doi.org/10.1159/000056789>
- Miron, E., Oldenkamp, R., Brown, J. M., Pinto, D. M. S., Xu, C. S., Faria, A. R., Shaban, H. A., Rhodes, J. D. P., Innocent, C., de Ornellas, S., Hess, H. F., Buckle, V., & Schermelleh, L. (2020). Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin. *Sci Adv*, 6(39). <https://doi.org/10.1126/sciadv.aba8811>
- Mohandas, T., Sparkes, R. S., & Shapiro, L. J. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science*, 211(4480), 393-396. <https://doi.org/10.1126/science.6164095>
- Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijsenhout, A., Pintacuda, G., Schermelleh, L., Nesterova, T. B., & Brockdorff, N. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep*, 12(4), 562-572. <https://doi.org/10.1016/j.celrep.2015.06.053>
- Monfort, A., Di Minin, G., Postlmayr, A., Freimann, R., Arieti, F., Thore, S., & Wutz, A. (2015). Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. *Cell Rep*, 12(4), 554-561. <https://doi.org/10.1016/j.celrep.2015.06.067>
- Moore, M. J., & Proudfoot, N. J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, 136(4), 688-700. <https://doi.org/10.1016/j.cell.2009.02.001>
- Mould, A. W., Pang, Z., Pakusch, M., Tonks, I. D., Stark, M., Carrie, D., Mukhopadhyay, P., Seidel, A., Ellis, J. J., Deakin, J., Wakefield, M. J., Krause, L., Blewitt, M. E., & Kay, G. F. (2013). Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin*, 6(1), 19. <https://doi.org/10.1186/1756-8935-6-19>
- Mueller, F., Stasevich, T. J., Mazza, D., & McNally, J. G. (2013). Quantifying transcription factor kinetics: at work or at play? *Crit Rev Biochem Mol Biol*, 48(5), 492-514. <https://doi.org/10.3109/10409238.2013.833891>
- Nativio, R., Wendt, K. S., Ito, Y., Huddleston, J. E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J. M., & Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet*, 5(11), e1000739. <https://doi.org/10.1371/journal.pgen.1000739>
- Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., & Avner, P. (2008). Molecular coupling of Xist regulation and pluripotency. *Science*, 321(5896), 1693-1695. <https://doi.org/10.1126/science.1160952>
- Nesterova, T. B., Slobodyanyuk, S. Y., Elisaphenko, E. A., Shevchenko, A. I., Johnston, C., Pavlova, M. E., Rogozin, I. B., Kolesnikov, N. N., Brockdorff, N., & Zakian, S. M. (2001). Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Res*, 11(5), 833-849. <https://doi.org/10.1101/gr.174901>

- Nesterova, T. B., Wei, G., Coker, H., Pintacuda, G., Bowness, J. S., Zhang, T., Almeida, M., Bloechl, B., Moindrot, B., Carter, E. J., Alvarez Rodrigo, I., Pan, Q., Bi, Y., Song, C. X., & Brockdorff, N. (2019). Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat Commun*, 10(1), 3129. <https://doi.org/10.1038/s41467-019-11171-3>
- Ng, K., Daigle, N., Bancaud, A., Ohhata, T., Humphreys, P., Walker, R., Ellenberg, J., & Wutz, A. (2011). A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. *Mol Biol Cell*, 22(14), 2634-2645. <https://doi.org/10.1091/mbc.E11-02-0146>
- Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V., & Laue, E. D. (2002). Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature*, 416(6876), 103-107. <https://doi.org/10.1038/nature722>
- Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Bluthgen, N., Dekker, J., & Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398), 381-385. <https://doi.org/10.1038/nature11049>
- Norris, D. P., Brockdorff, N., & Rastan, S. (1991). Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. *Mamm Genome*, 1(2), 78-83. <https://doi.org/10.1007/BF02443782>
- Nozawa, R. S., Nagao, K., Igami, K. T., Shibata, S., Shirai, N., Nozaki, N., Sado, T., Kimura, H., & Obuse, C. (2013). Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBIX1 pathway. *Nat Struct Mol Biol*, 20(5), 566-573. <https://doi.org/10.1038/nsmb.2532>
- Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D., & Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science*, 303(5658), 644-649. <https://doi.org/10.1126/science.1092727>
- Padeken, J., Methot, S. P., & Gasser, S. M. (2022). Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat Rev Mol Cell Biol*, 23(9), 623-640. <https://doi.org/10.1038/s41580-022-00483-w>
- Panaretou, B., Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1998). ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo. *EMBO J*, 17(16), 4829-4836. <https://doi.org/10.1093/emboj/17.16.4829>
- Pandya-Jones, A., Markaki, Y., Serizay, J., Chitiashvili, T., Mancía Leon, W. R., Damianov, A., Chronis, C., Papp, B., Chen, C. K., McKee, R., Wang, X. J., Chau, A., Sabri, S., Leonhardt, H., Zheng, S., Guttman, M., Black, D. L., & Plath, K. (2020). A protein assembly mediates Xist localization and gene silencing. *Nature*, 587(7832), 145-151. <https://doi.org/10.1038/s41586-020-2703-0>
- Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B. S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A. G., & Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, 132(3), 422-433. <https://doi.org/10.1016/j.cell.2008.01.011>

- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am J Hum Genet*, 31(2), 106-115. <https://www.ncbi.nlm.nih.gov/pubmed/377956>
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., & Jaffrey, S. R. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature*, 537(7620), 369-373. <https://doi.org/10.1038/nature19342>
- Paweletz, N. (2001). Walther Flemming: pioneer of mitosis research. *Nat Rev Mol Cell Biol*, 2(1), 72-75. <https://doi.org/10.1038/35048077>
- Pedersen, L. C., Inoue, K., Kim, S., Perera, L., & Shaw, N. D. (2019). A ubiquitin-like domain is required for stabilizing the N-terminal ATPase module of human SMCHD1. *Commun Biol*, 2, 255. <https://doi.org/10.1038/s42003-019-0499-y>
- Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., & Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature*, 379(6561), 131-137. <https://doi.org/10.1038/379131a0>
- Pepenella, S., Murphy, K. J., & Hayes, J. J. (2014). Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. *Chromosoma*, 123(1-2), 3-13. <https://doi.org/10.1007/s00412-013-0435-8>
- Petela, N. J., Gligoris, T. G., Metson, J., Lee, B. G., Voulgaris, M., Hu, B., Kikuchi, S., Chapard, C., Chen, W., Rajendra, E., Srinivisan, M., Yu, H., Lowe, J., & Nasmyth, K. A. (2018). Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. *Mol Cell*, 70(6), 1134-1148 e1137. <https://doi.org/10.1016/j.molcel.2018.05.022>
- Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B. A., Solcan, N., Cerase, A., Castello, A., Mohammed, S., Moindrot, B., Nesterova, T. B., & Brockdorff, N. (2017). hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol Cell*, 68(5), 955-969 e910. <https://doi.org/10.1016/j.molcel.2017.11.013>
- Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B., & Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science*, 300(5616), 131-135. <https://doi.org/10.1126/science.1084274>
- Poonperm, R., Ichihara, S., Miura, H., Tanigawa, A., Nagao, K., Obuse, C., Sado, T., & Hiratani, I. (2023). Replication dynamics identifies the folding principles of the inactive X chromosome. *Nat Struct Mol Biol*, 30(8), 1224-1237. <https://doi.org/10.1038/s41594-023-01052-1>
- Presman, D. M., Ball, D. A., Paakinaho, V., Grimm, J. B., Lavis, L. D., Karpova, T. S., & Hager, G. L. (2017). Quantifying transcription factor binding dynamics at the single-molecule level in live cells. *Methods*, 123, 76-88. <https://doi.org/10.1016/j.ymeth.2017.03.014>
- Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell*, 90(1), 65-75. [https://doi.org/10.1016/s0092-8674\(00\)80314-1](https://doi.org/10.1016/s0092-8674(00)80314-1)
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, 8(11), 2281-2308. <https://doi.org/10.1038/nprot.2013.143>

- Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., & Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7), 1665-1680. <https://doi.org/10.1016/j.cell.2014.11.021>
- Rastan, S., & Robertson, E. J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J Embryol Exp Morphol*, 90, 379-388. <https://www.ncbi.nlm.nih.gov/pubmed/3834036>
- Rens, W., Wallduck, M. S., Lovell, F. L., Ferguson-Smith, M. A., & Ferguson-Smith, A. C. (2010). Epigenetic modifications on X chromosomes in marsupial and monotreme mammals and implications for evolution of dosage compensation. *Proc Natl Acad Sci U S A*, 107(41), 17657-17662. <https://doi.org/10.1073/pnas.0910322107>
- Rhodes, J., Mazza, D., Nasmyth, K., & Uphoff, S. (2017). Scc2/Nipbl hops between chromosomal cohesin rings after loading. *Elife*, 6. <https://doi.org/10.7554/eLife.30000>
- Ridings-Figueroa, R., Stewart, E. R., Nesterova, T. B., Coker, H., Pintacuda, G., Godwin, J., Wilson, R., Haslam, A., Lilley, F., Ruigrok, R., Bageghni, S. A., Albadrani, G., Mansfield, W., Roulson, J. A., Brockdorff, N., Ainscough, J. F. X., & Coverley, D. (2017). The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes Dev*, 31(9), 876-888. <https://doi.org/10.1101/gad.295907.117>
- Rodermund, L., Coker, H., Oldenkamp, R., Wei, G., Bowness, J., Rajkumar, B., Nesterova, T., Susano Pinto, D. M., Schermelleh, L., & Brockdorff, N. (2021). Time-resolved structured illumination microscopy reveals key principles of Xist RNA spreading. *Science*, 372(6547). <https://doi.org/10.1126/science.abe7500>
- Sacconi, S., Lemmers, R. J., Balog, J., van der Vliet, P. J., Lahaut, P., van Nieuwenhuizen, M. P., Straasheijm, K. R., Debipersad, R. D., Vos-Versteeg, M., Salvati, L., Casarin, A., Pegoraro, E., Tawil, R., Bakker, E., Tapscott, S. J., Desnuelle, C., & van der Maarel, S. M. (2013). The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet*, 93(4), 744-751. <https://doi.org/10.1016/j.ajhg.2013.08.004>
- Sado, T., Hoki, Y., & Sasaki, H. (2006). Tsix defective in splicing is competent to establish Xist silencing. *Development*, 133(24), 4925-4931. <https://doi.org/10.1242/dev.02670>
- Sado, T., Okano, M., Li, E., & Sasaki, H. (2004). De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development*, 131(5), 975-982. <https://doi.org/10.1242/dev.00995>
- Sakaguchi, T., Hasegawa, Y., Brockdorff, N., Tsutsui, K., Tsutsui, K. M., Sado, T., & Nakagawa, S. (2016). Control of Chromosomal Localization of Xist by hnRNP U Family Molecules. *Dev Cell*, 39(1), 11-12. <https://doi.org/10.1016/j.devcel.2016.09.022>
- Sakakibara, Y., Nagao, K., Blewitt, M., Sasaki, H., Obuse, C., & Sado, T. (2018). Role of SmcHD1 in establishment of epigenetic states required for the maintenance of the X-inactivated state in mice. *Development*, 145(18). <https://doi.org/10.1242/dev.166462>

- Saksouk, N., Barth, T. K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M. E., Imhof, A., Dejardin, J., & Simboeck, E. (2014). Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol Cell*, 56(4), 580-594. <https://doi.org/10.1016/j.molcel.2014.10.001>
- Sanborn, A. L., Rao, S. S., Huang, S. C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K. P., Gnirke, A., Melnikov, A., McKenna, D., Stamenova, E. K., Lander, E. S., & Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A*, 112(47), E6456-6465. <https://doi.org/10.1073/pnas.1518552112>
- Schmiedeberg, L., Skene, P., Deaton, A., & Bird, A. (2009). A temporal threshold for formaldehyde crosslinking and fixation. *PLoS One*, 4(2), e4636. <https://doi.org/10.1371/journal.pone.0004636>
- Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., & Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J*, 21(5), 1121-1131. <https://doi.org/10.1093/emboj/21.5.1121>
- Schulz, E. G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Bluthgen, N., & Heard, E. (2014). The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell*, 14(2), 203-216. <https://doi.org/10.1016/j.stem.2013.11.022>
- Shaw, N. D., Brand, H., Kupchinsky, Z. A., Bengani, H., Plummer, L., Jones, T. I., Erdin, S., Williamson, K. A., Rainger, J., Stortchevoi, A., Samocha, K., Currall, B. B., Dunican, D. S., Collins, R. L., Willer, J. R., Lek, A., Lek, M., Nassan, M., Pereira, S., . . . Talkowski, M. E. (2017). SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet*, 49(2), 238-248. <https://doi.org/10.1038/ng.3743>
- Shibata, S., & Lee, J. T. (2004). Tsix transcription- versus RNA-based mechanisms in Xist repression and epigenetic choice. *Curr Biol*, 14(19), 1747-1754. <https://doi.org/10.1016/j.cub.2004.09.053>
- Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P., & Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell*, 4(4), 481-495. [https://doi.org/10.1016/s1534-5807\(03\)00068-6](https://doi.org/10.1016/s1534-5807(03)00068-6)
- Simon, M. D., Pinter, S. F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S. K., Kesner, B. A., Maier, V. K., Kingston, R. E., & Lee, J. T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature*, 504(7480), 465-469. <https://doi.org/10.1038/nature12719>
- Smeets, D., Markaki, Y., Schmid, V. J., Kraus, F., Tattermusch, A., Cerase, A., Sterr, M., Fiedler, S., Demmerle, J., Popken, J., Leonhardt, H., Brockdorff, N., Cremer, T., Schermelleh, L., & Cremer, M. (2014). Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its

- active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin*, 7, 8. <https://doi.org/10.1186/1756-8935-7-8>
- Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J., Zhu, Y., Kaaij, L. J., van Ijcken, W., Gribnau, J., Heard, E., & de Laat, W. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev*, 25(13), 1371-1383. <https://doi.org/10.1101/gad.633311>
- Sun, B. K., Deaton, A. M., & Lee, J. T. (2006). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell*, 21(5), 617-628. <https://doi.org/10.1016/j.molcel.2006.01.028>
- Sunwoo, H., Colognori, D., Froberg, J. E., Jeon, Y., & Lee, J. T. (2017). Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). *Proc Natl Acad Sci U S A*, 114(40), 10654-10659. <https://doi.org/10.1073/pnas.1711206114>
- Sunwoo, H., Wu, J. Y., & Lee, J. T. (2015). The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc Natl Acad Sci U S A*, 112(31), E4216-4225. <https://doi.org/10.1073/pnas.1503690112>
- Szczurek, A. T., Dimitrova, E., Kelley, J. R., Blackledge, N. P., & Klose, R. J. (2024). The Polycomb system sustains promoters in a deep OFF state by limiting pre-initiation complex formation to counteract transcription. *Nat Cell Biol.* <https://doi.org/10.1038/s41556-024-01493-w>
- Takagi, N., & Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, 256(5519), 640-642. <https://doi.org/10.1038/256640a0>
- Tapia Del Fierro, A., den Hamer, B., Benetti, N., Jansz, N., Chen, K., Beck, T., Vanyai, H., Gurzau, A. D., Daxinger, L., Xue, S., Ly, T. T. N., Wanigasuriya, I., Iminoff, M., Breslin, K., Oey, H., Krom, Y. D., van der Hoorn, D., Bouwman, L. F., Johanson, T. M., . . . Blewitt, M. E. (2023). SMCHD1 has separable roles in chromatin architecture and gene silencing that could be targeted in disease. *Nat Commun*, 14(1), 5466. <https://doi.org/10.1038/s41467-023-40992-6>
- Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., & Brockdorff, N. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell*, 148(4), 664-678. <https://doi.org/10.1016/j.cell.2011.12.029>
- Teller, K., Illner, D., Thamm, S., Casas-Delucchi, C. S., Versteeg, R., Indemans, M., Cremer, T., & Cremer, M. (2011). A top-down analysis of Xa- and Xi-territories reveals differences of higher order structure at ≥ 20 Mb genomic length scales. *Nucleus*, 2(5), 465-477. <https://doi.org/10.4161/nucl.2.5.17862>
- Tian, D., Sun, S., & Lee, J. T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell*, 143(3), 390-403. <https://doi.org/10.1016/j.cell.2010.09.049>
- Tokunaga, M., Imamoto, N., & Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods*, 5(2), 159-161. <https://doi.org/10.1038/nmeth1171>

- Tonkin, E. T., Wang, T. J., Lisgo, S., Bamshad, M. J., & Strachan, T. (2004). NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet*, 36(6), 636-641. <https://doi.org/10.1038/ng1363>
- Trojer, P., & Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell*, 28(1), 1-13. <https://doi.org/10.1016/j.molcel.2007.09.011>
- Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., & Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J*, 13(16), 3822-3831. <https://doi.org/10.1002/j.1460-2075.1994.tb06693.x>
- Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D. J., & Kapanidis, A. N. (2013). Single-molecule DNA repair in live bacteria. *Proc Natl Acad Sci U S A*, 110(20), 8063-8068. <https://doi.org/10.1073/pnas.1301804110>
- Uphoff, S., Sherratt, D. J., & Kapanidis, A. N. (2014). Visualizing protein-DNA interactions in live bacterial cells using photoactivated single-molecule tracking. *J Vis Exp*(85). <https://doi.org/10.3791/51177>
- van Bommel, J. G., Galupa, R., Gard, C., Servant, N., Picard, C., Davies, J., Szempruch, A. J., Zhan, Y., Zylicz, J. J., Nora, E. P., Lameiras, S., de Wit, E., Gentien, D., Baulande, S., Giorgetti, L., Guttman, M., Hughes, J. R., Higgs, D. R., Gribnau, J., & Heard, E. (2019). The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat Genet*, 51(6), 1024-1034. <https://doi.org/10.1038/s41588-019-0412-0>
- Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K. K., Olsen, J. V., Hyman, A. A., Stunnenberg, H. G., & Mann, M. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, 142(6), 967-980. <https://doi.org/10.1016/j.cell.2010.08.020>
- Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M., & Timmers, H. T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, 131(1), 58-69. <https://doi.org/10.1016/j.cell.2007.08.016>
- Waddington, C. H. (1942). The epigenotype. 1942. *Int J Epidemiol*, 41(1), 10-13. <https://doi.org/10.1093/ije/dyr184>
- Wang, C. Y., Brand, H., Shaw, N. D., Talkowski, M. E., & Lee, J. T. (2019). Role of the Chromosome Architectural Factor SMCHD1 in X-Chromosome Inactivation, Gene Regulation, and Disease in Humans. *Genetics*, 213(2), 685-703. <https://doi.org/10.1534/genetics.119.302600>
- Wang, C. Y., Colognori, D., Sunwoo, H., Wang, D., & Lee, J. T. (2019). PRC1 collaborates with SMCHD1 to fold the X-chromosome and spread Xist RNA between chromosome compartments. *Nat Commun*, 10(1), 2950. <https://doi.org/10.1038/s41467-019-10755-3>
- Wang, C. Y., Jegu, T., Chu, H. P., Oh, H. J., & Lee, J. T. (2018). SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. *Cell*, 174(2), 406-421 e425. <https://doi.org/10.1016/j.cell.2018.05.007>

- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., & Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature*, 431(7010), 873-878. <https://doi.org/10.1038/nature02985>
- Wang, Z., Song, J., Milne, T. A., Wang, G. G., Li, H., Allis, C. D., & Patel, D. J. (2010). Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CYP33 and HDAC-mediated repression. *Cell*, 141(7), 1183-1194. <https://doi.org/10.1016/j.cell.2010.05.016>
- Wanigasuriya, I., Kinkel, S. A., Beck, T., Roper, E. A., Breslin, K., Lee, H. J., Keniry, A., Ritchie, M. E., Blewitt, M. E., & Gouil, Q. (2022). Maternal SMCHD1 controls both imprinted Xist expression and imprinted X chromosome inactivation. *Epigenetics Chromatin*, 15(1), 26. <https://doi.org/10.1186/s13072-022-00458-3>
- Wei, G. (2024). RNA m6A modification, signals for degradation or stabilisation? *Biochem Soc Trans*, 52(2), 707-717. <https://doi.org/10.1042/BST20230574>
- Wei, H., Ruthenburg, A. J., Bechis, S. K., & Verdine, G. L. (2005). Nucleotide-dependent domain movement in the ATPase domain of a human type IIA DNA topoisomerase. *J Biol Chem*, 280(44), 37041-37047. <https://doi.org/10.1074/jbc.M506520200>
- Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., & Peters, J. M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, 451(7180), 796-801. <https://doi.org/10.1038/nature06634>
- Wozniak, G. G., & Strahl, B. D. (2014). Hitting the 'mark': interpreting lysine methylation in the context of active transcription. *Biochim Biophys Acta*, 1839(12), 1353-1361. <https://doi.org/10.1016/j.bbagra.2014.03.002>
- Wutz, A., Rasmussen, T. P., & Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet*, 30(2), 167-174. <https://doi.org/10.1038/ng820>
- Yamada, N., Hasegawa, Y., Yue, M., Hamada, T., Nakagawa, S., & Ogawa, Y. (2015). Xist Exon 7 Contributes to the Stable Localization of Xist RNA on the Inactive X-Chromosome. *PLoS Genet*, 11(8), e1005430. <https://doi.org/10.1371/journal.pgen.1005430>
- Young, A. N., Perlas, E., Ruiz-Blanes, N., Hierholzer, A., Pomella, N., Martin-Martin, B., Liverziani, A., Jachowicz, J. W., Giannakouros, T., & Cerase, A. (2021). Deletion of LBR N-terminal domains recapitulates Pelger-Huet anomaly phenotypes in mouse without disrupting X chromosome inactivation. *Commun Biol*, 4(1), 478. <https://doi.org/10.1038/s42003-021-01944-2>
- Zaret, K. S., & Carroll, J. S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*, 25(21), 2227-2241. <https://doi.org/10.1101/gad.176826.111>
- Zylicz, J. J., Bousard, A., Zumer, K., Dossin, F., Mohammad, E., da Rocha, S. T., Schwalb, B., Syx, L., Dingli, F., Loew, D., Cramer, P., & Heard, E. (2019). The Implication of Early Chromatin Changes in X Chromosome Inactivation. *Cell*, 176(1-2), 182-197 e123. <https://doi.org/10.1016/j.cell.2018.11.041>
- Zylicz, J. J., & Heard, E. (2020). Molecular Mechanisms of Facultative Heterochromatin Formation: An X-Chromosome Perspective. *Annu Rev Biochem*, 89, 255-282. <https://doi.org/10.1146/annurev-biochem-062917-012655>

